The Hogan RAW Assessment: Technical Manual v3

Reece Akhtar, Brandon Ferrell, Dave Winsborough, & Uri Ort

Contents

Figures

Tables

1. Conceptual Background

A shift is underway in both the form of employment and the nature of the contract between employers and employees, seen most easily in the rise of 'zero hours' contracts, the 'gig' economy and the commodification of work via online platforms like Uber or Task Rabbit (De Stefano, 2015; Huws & Joyce, 2016). These disruptive technologies, combined with the emergence of artificial intelligence and robotics, and a long term decline in labor market flexibility, has observers and job seekers worried that the nature of employment is permanently changed (Molloy, Smith, Trezzi, & Wozniak, 2016). Further evidence suggests companies are hiring from the outside rather than growing their own talent and they are investing less in training (Cappelli, 2015). Although firms may complain about the difficulty in finding talent (Cappelli, 2015), the real issue is the poor job done in matching people to workplaces and careers. The global economy is characterized by increasingly diverse employment patterns, rising risk and uncertainty for many workers, as well as rapidly growing inequalities in rewards (Fuller & Stecy-Hildebrandt, 2014).

The illiquidity of the labor market offers both business opportunities and the chance to help people avoid unemployment, especially the young, who are disproportionally impacted by economic cycles (Pallais, 2014). The Pulitzer Prize winning author J.R. Moeringher once described "the sadness and terror of joblessness" and called unemployment "bad for the soul" (Lashinsky, 2012). Certainly the evidence is that unemployment, under-employment and losing work creates a host of negative life outcomes, including long-term income losses, reduction in psychological and physical well-being, loss of psychosocial assets, social withdrawal, increased suicide risk, family discord, and lower levels of children's attainment and health (Brand, 2015; Milner, Page, & LaMontagne, 2013; Roelfs, Shor, Davidson, & Schwartz, 2011). Work has benefits that extend beyond the purely economic (Fuller & Stecy-Hildebrandt, 2014); workers derive structure, social connection, meaning, stimulation and even enjoyment from their labors. Helping people into better careers and jobs is therefore a worthwhile endeavor.

Although there are many causes of unemployment, typical analyses focus on cyclical versus structural economic factors. Cyclical unemployment occurs when demand for goods and services falls, which results in fewer jobs. Structural unemployment refers to a mismatch between the qualifications employers seek and the skills of the available labor force. For example, evidence suggests that both factors explain the current global trend toward increasing underemployment and that structural unemployment has been on the rise in the United States for some time (Levine, 2013).

While economic factors affect employment status, psychological factors associated with employability are also known to be important (Fugate, Kinicki, & Ashforth, 2004; Hogan, Chamorro-Premuzic & Kaiser, 2013). For these reasons, finding and sustaining employment is one of the most significant skills any adult can acquire. That is why we designed this tool as a means of understanding the likelihood of an individual succeeding or struggling in gaining employment, and to help test takers focus on jobs at which they are more likely to do well.

1.1. Employment Success

Organizational psychology has had a lot to say about what it takes to do well in the world of work. We briefly review research showing that cognitive ability, educational attainment and personality interact to make some people do better at work and get ahead in their careers.

The literature clearly shows that cognitive ability and personality both predict educational performance, which is associated with work success and job performance. For example a large evaluation of the links between personality and income, occupational prestige, long-term unemployment, and occupational stability, concluded that personality scales predict work-related outcomes and that the typical effect size of personality is similar to the impact of childhood socioeconomic status and IQ (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). Similarly, Swedish economists examined the relationship of cognitive and non-cognitive skills on labor market outcomes (Lindqvist, & Vestman, 2011) and found that personality factors (e.g., responsibility, independence, persistence, emotional stability, and social skill) were, in combination, better predictors of wages, employment status, and annual earnings. Consistent with these conclusions, longitudinal research showed that personality and IQ assessed in childhood each predicted occupational status in adulthood, yielding a multiple correlation of .64 (Judge, Higgins, Thoresen, & Barrick, 1999).

Our point is that while macro-level economic factors impact success and failure at work, so do individual factors. A large meta-analysis of research on career success (Ng, Eby, Sorensen, & Feldman, 2005) found that promotion, salary and career satisfaction was dependent on two key elements: a) working hard and displaying relevant abilities and skills, and b) receiving sponsorship. The latter point is important as the conventional wisdom of applied psychology maintains that employers should be most interested in the degree to which applicants possess smarts and useful personality characteristics (Baruch & Bozionelos, 2011). Yet career success also depends on factors outside the control of individual actors; career success is affected by the political structure of organizations. Changes in leadership are often accompanied by other staffing changes, and alliances can determine who gets which job — or no job at all. The study by Ng et al. (2011) concluded that sponsorship by powerful others was dependent on employees getting along with their managers, being positive and self-promoting, or in other words, their social skills. A well-designed German study found that job performance ratings are strongly predicted by the motive to get along with others and the motive to achieve status and power (Blickle et al., 2011). In a Dutch study of 315 entry level employees, occupational competence predicted income but the best predictor of employability and career success was the ability to maintain socially desirable performance at work (Van der Heijde & Van der Heijden, 2006).

1.2. What Employers Want

If career success depends in large measure with the employee's ability to meet employer expectations then it is important to understand what employers want in their employees — after all, hiring organizations ultimately decide who is employed. Surprisingly, the literature on this is more sparse (Baruch & Bozionelos, 2011). Broadly, the modern workplace requires workers to have cognitive and affective skills, which employers are increasingly keen to seek straight out of the box (Cappelli & Keller, 2014; Koenig, 2011). Below, we survey what can be said about what employers want.

The U.S. Department of Labor undertook a large scale study of what employers look for in prospective employees (The Secretary's Commission on Achieving Necessary Skills, 1991). By surveying business owners, union officials, public employees, managers, and private-sector workers to determine the performance demands of modern employment, they identified five critical competencies sought by employers: (a) ability to identify and allocate resources; (b) ability to work with others; (c) ability to acquire and use information; (d) ability to understand complex interrelationships; and (e) ability to work with a variety of technologies.

A comprehensive content analysis of employment adverts from the United States offers another view of what employers look for (Hogan & Brinkmeyer, 1994). The content of over 6000 job advertisements was analyzed, showing that nearly half of all jobs required good interpersonal skills. Social skill was vital in 71% of the jobs involving client contact, 78% of the jobs requiring coworker interaction, and 83% of managerial roles. The conclusion is clear: from the employers' perspective, the single-most important characteristic determining employability is interpersonal skill or social competence.

Looking at this matter from the perspective of what employers *don't want* yields further insights that underpin the previous discussion. A large-scale survey found that employers complain most about three kinds of worker deficiencies: (a) poor problem solving, (b) poor personal management, and (c) poor interpersonal skills (Bureau of National Affairs, 1988).

In summary, the literature on what employers want in job candidates highlights the importance of problem solving, social skills and being a motivated colleague. While these may seem to be common sense findings, the fact remains that people vary on these dimensions, making some more attractive to employers, and others less so.

1.3. An Integrative Model of Employability

We characterize employability as the ability to gain and maintain a job in a formal organization (Fugate et al., 2004). The deduction from the forgoing discussion is that employability is an attribution employers make about job seekers, which is based typically on a resume scan or interview — neither of which are very effective predictors of job performance (Chamorro-Premuzic & Furnham, 2010). The essential question is therefore: What determines whether a person will be perceived as having the potential to contribute positively to an organization?

Job performance is primarily defined by supervisors' ratings. In very general terms, supervisors want employees who are likeable. In addition, they favor employees who seem to learn quickly and show good judgment — and this helps explain the consistent correlations between cognitive ability and job performance (Kuncel, Ones, & Sackett, 2010). Supervisors also like employees who seem compliant, obedient, and conforming — and this helps explain the consistent correlations between measures of conscientiousness and job performance (Chamorro-Premuzic & Furnham, 2010). These observations also account for the positive manifold between personality, cognitive ability, educational performance, and job performance.

Our answer to the question of what determines a person's level of employability will be a function of (a) interpersonal skill (Lievens & Sackett, 2012) and perceived compatibility with the values of the organization, team, or management (Edwards & Cable, 2009); (b) ability, know-how , and expertise (Schmidt, Hunter, & Outerbridge, 1986); and (c) ambition, drive, and work ethic (Mikulincer, Shaver, Simpson, & Dovidio, 2015). This model can be more simply summarized using the acronym RAW (Hogan, Chamorro-Premuzic & Kaiser, 2013): a person is more employable if they are (a) rewarding to deal with - Rewarding; (b) capable of learning and performing the job - Able; and (c) driven and hardworking – Willing (see Figure 1).

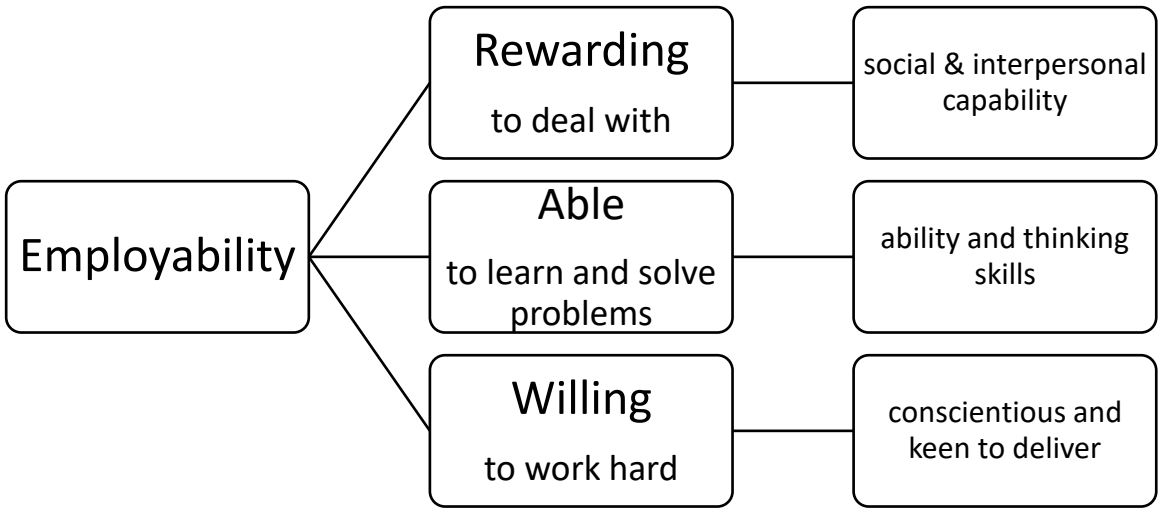

Figure 1: The Determinants of Employability

Importantly, this model is compensatory. That is, employees with only average ability can still succeed by being rewarding and productive; bright people with limited social skills may succeed by being very productive; those who are charming, bright, and lazy will succeed as they always do. But candidates who ''tick all three boxes'' can be expected to enjoy higher levels of career success; candidates who are strong in two of the three areas should do well; those who are strong in only one can expect occasional unemployment; those who are weak in all three may find the world of work hard to navigate and a harsh environment.

1.4. Careers and Vocational Interests

Career success is proximally determined by the ability to gain employment in the first place – to get a foot on the career ladder. More distally, success depends on how it is defined. Salary, promotion, and individual career satisfaction represent conceptually distinct aspects of career success (Ng et al., 2005). A person may earn a good wage or salary and yet find themselves unhappy and unsatisfied at work; others may find work is a calling, and that objective indicators of success are irrelevant (Hall & Chandler, 2004). For that reason, models of job and career fit have emphasized the role of vocational interests and values when predicting satisfaction and success (Ehrhart & Makransky, 2007; Holland, 1997; Stoll et al., 2016). Importantly, research demonstrating the validity of interests and vocational fit

in predicting relevant work behaviors, including job knowledge, job performance (Van Iddekinge, Roth, Putka, & Lanivich, 2011).

Holland's (1997) taxonomy of vocational preferences has been the gold standard for over 40 years. It provides an integrative framework that proposes that a set of six interest-based categories—Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC) can be used to understand work interests. Thus, all careers can be classified by preferences for certain types of tasks, environments and interactional styles:

1. *Realistic*: People who have athletic or technical interests, prefer to work with objects, machines, tools, plants and animals.
2. *Investigative*: People who like to observe, learn, investigate, analyse, evaluate or solve problems.
3. *Artistic*: People who have artistic, aesthetic, and creative interests and prefer to work in unstructured environments exercising their imagination and creativity.
4. *Social*: People who like to work with others to enlighten, inform, help, train, or cure them.
5. *Enterprising*: People who like to influence, persuade, manage and lead others to achieve organizational or economic success.
6. *Conventional*: People who like to work with data, carrying out tasks that require following procedures and paying attention to detail.

Decades of research support the idea that people whose personality fits their work environments will be happier and more satisfied. It's also been shown that when switching jobs, individuals move to environments and work that tends to suit their personality (Nauta, 2010; Wille, De Fruyt, Dingemanse, & Vergauwe, 2015).

Because both jobs and people can be classified with RIASEC it is possible to quickly determine any individual's career interests. Importantly, since RIASEC is used by the world's most comprehensive listing of jobs, the US Department of Labor's O*Net, it becomes possible to match people to real careers and jobs. Using other data, like education level or aptitude, these matches can be refined to indicate jobs that fit their likely qualifications and aptitude. It is for these reasons that the RIASEC model complements any assessment of employability. It can be argued that putting the right people (those individuals with high employability) in the right jobs (as indicated by their RIASEC preferences) will increase their potential to perform and be motivated at work.

2. The Hogan RAW Assessment

Given the economic, societal, and individual importance of employability (defined as the ability to gain and maintain a job—Fugate, et al., 2004), it would be useful to have a valid method for predicting peoples' potential for employability, something that could be used to inform hiring and placement decisions. As described above, we identify three dimensions contribute to peoples' employability:

1. *Rewarding*: Being able to get along with colleagues, having good interpersonal skills.
2. *Able*: Being able to learn the job, developing expertise and relevant knowledge.
3. *Willing*: Being able to work hard, and work to a high standard of performance.

Individuals with high scores on all three dimensions are highly employable; individuals with low scores on all three dimensions will have difficulty gaining and/or maintaining employment.

The following chapter describes the development of our RAW assessment. Subsequent chapters provide information regarding its psychometric reliability and validity, and demonstrate that the assessment has no adverse impact when used in selection contexts.

2.1. The Rewarding & Willing Scales

To determine the item content for the two scales, we asked a group of Subject Matter Experts (SMEs) with advanced degrees in Industrial-Organizational Psychology to review Saucier's (1994) list of "mini-markers" (40 adjective markers for each of the Big Five dimensions found in phenotypic personality description) and identify those that best describe Rewarding and Willing, as defined by Hogan, et al. (2013). The SMEs converged nicely in their choices of mini-markers describing Rewarding and Willing individuals.

To identify items to assess the two dimensions, we used archival data and correlated items from the Hogan Personality Inventory (Hogan & Hogan, 2007) and the Hogan Development Survey (Hogan & Hogan, 2009) with the mini-markers. Items that correlated with the Rewarding markers or the Willing markers were retained for further investigation. We used factor analysis to reduce the number of items; the analysis yielded a two-factor solution (Factor 1 = Rewarding and Factor 2 = Willing) with no cross-loading items and we retained 12 items, six items for the Rewarding dimension and six for the Willing dimension. The two six-item scales are uncorrelated. Participants respond to the items using a four point Likert scale (Strongly Agree, Agree, Disagree, Strongly Disagree), and each scale is scored by summing the responses.

Individuals with high Rewarding scores can be expected to be relaxed, easygoing and confident, whereas people with low Rewarding scores are likely to be irritable, emotionally volatile and lacking self-confidence. Individuals with high Willing scores will hold themselves to high standards, set lofty goals, and be highly disciplined. On the other hand, individuals with low Willing scores should seem disorganized, unreliable, and unconcerned about the quality of their work.

2.2. The Able Game

To assess Able, we partnered with MindX, a technology firm that creates game-based assessments. Respondents find gamified ability assessments enjoyable (Lumsden, Edwards, Lawrence, Coyle, & Munafò, 2016).

Specifically, we developed two cognitive ability games: "*Shape Dance*" and "*Disco Numbers*". Shape Dance is an abstract reasoning and mental rotation task. Players are shown two or more grids and must decide whether the grids are the same. The game becomes more difficult as players are given grids of different sizes, complexity, and orientations.

Disco Numbers is a numerical reasoning and working memory game. Participants are shown a sequence of numbers which they memorize and compute the sum of the sequence. This game becomes more difficult as the sequences get longer and require more difficult mathematics.

The participants' goal in both games is to provide as many correct responses as possible; the maximum score in both games is 30. The average score across both games is our index of the Able dimension of employability. For information regarding the games' scoring and validity, see **Chapter 3.3.3**. Finally, each game takes three minutes to complete. Participants also complete practice rounds to familiarize themselves with the games' design and mechanics. Overall, the game session lasts around 10 minutes. **Figure 2** are examples of both games at their easiest level of difficulty.

When interpreting Able scores, individuals with higher scores easily solve problems, learn new skills, and mentally manipulate information. Conversely, individuals with low scores are likely to struggle to understand abstract concepts, to solve problems and acquire expertise quickly.

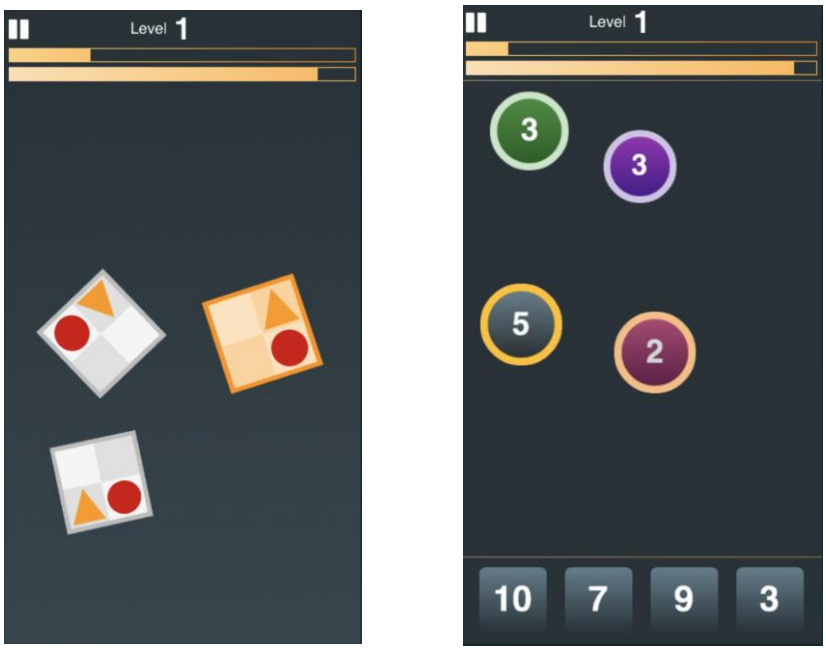

Figure 2: Example of the Shape Dance (*Left*) and Disco Number (*Right*) Games.

3. The Psychometric Properties of the Hogan RAW Assessment

3.1. Descriptive & Reliability Statistics

Table 1 presents descriptive statistics for the dimensions of the RAW assessment, including minimum and maximum observed scores, mean scores, standard deviations, skewness, kurtosis, and internal consistency, statistics[1].

To examine the psychometric properties of the dimensions, we obtained data from a sample of 2,452 individuals. On average, participants were 36.27 years old (*SD* = 11.88), and 45.99% of the sample were male while 49.88% were female (4.13% of participants did not indicate their sex). Furthermore, 53.24% were in full-time employment, 13.37% were self-employed, 11.55% were in part-time employment, 5.27% were students and 9.60% were unemployed.

Table 1: Classical Scale Statistics for the Hogan RAW Assessment

Dimension	Min	Max	*M*	*SD*	Skewness	Kurtosis	α
Rewarding	6	24	15.68	3.45	-.17	.05	.78
Willing	7	24	17.57	2.88	-.06	.06	.69
Able	.00	.80	.49	.13	-.66	.70	.46

Note: Total N = 2,452; Rewarding N = 2,318; Willing N = 2,312; Able = 498; Min = Minimum score; Max = Maximum score; *M* = Mean; *SD* = Standard Deviation; α = Cronbach's alpha.

The Rewarding and Willing scales have a maximum score of 24; Able is scored using the average of correct responses across two games and has a maximum score of 1.00. As shown in **Table 1**, the observed scores for the Rewarding and Willing dimensions cover the entire range of possible scores, while observed Able scores do not.

These data indicate that the assessment has acceptable technical properties.

[1] Skewness refers to departure from symmetry in a distribution of scores. When a distribution is normal and symmetrical, skewness values are around zero. Positive skewness values indicate that most scores fall at the bottom end of a distribution, and negative skewness values indicate that most scores fall on the top end of a distribution. Skewness values greater than +1.0 or less than -1.0 generally indicate a significant departure from symmetry.

Kurtosis refers to how peaked or flat a score distribution is relative to the normal distribution. When scores are normally distributed, kurtosis values are around zero and we refer to them as mesokurtic. When the distribution is sharper than the normal distribution, kurtosis values are positive and we refer to them as leptokurtic. When the distribution is broader than the normal distribution, kurtosis values are negative and we refer to them as platykurtic. Kurtosis values of more than twice the standard error indicate a significant departure from the normal distribution.

Internal consistency is a form of reliability estimating how well variables—in this case the RAW dimensions—estimate a common attribute. When scales measure the same construct, internal consistency is high. When scales measure different constructs, it is low. Internal consistency is most commonly estimated using Cronbach's alpha. Alpha estimates higher than .60 are acceptable and those higher than .70 are considered good.

Although the Rewarding and Willing scales show acceptable levels of internal consistency, the internal consistency of the Able measure is low. This is not a significant problem for two reasons. First, a principal components analysis of the proportion of wins to losses across the two games reveals a single latent factor. Second, this score is strongly correlated with traditional measures of cognitive ability (for more information, see **Chapter 3.3.3.2**).

3.2. Distribution of RAW Scores

We transformed participant's observed scores on each of the three dimensions into percentiles, so that cutoff scores could be computed across four levels: Low, Below Average, Above Average, and High. Using these ranges, we examined score distributions for all dimensions using the same sample previously described. **Table 2** presents these results and lists specific cutoff scores for each of the three dimensions.

Table 2: Distribution of Hogan RAW Assessment Scores Across Levels

Dimension	Low (0 – 25%)	Moderately Low (26%-50%)	Moderately High (51% - 75%)	High (76% - 100%)
Rewarding	6 - 13	14 - 16	17 - 18	19 - 30
Willing	6 - 16	17 - 18	19	20 - 30
Able	.00 - .40	.41 - .49	.50 - .57	.58 - 1.00

Note: Total N = 2,452; Rewarding N = 2,318; Willing N = 2,317; Able = 498. Percentiles and interpretative levels are calculated using scores that are rounded to two decimal places.

3.3. The RAW Dimensions

The following section contains information regarding the construction of the three RAW scales. Specifically, results regarding descriptive statistics, intercorrelations, and scoring methods provide a better understanding of the psychometric properties of the RAW assessment.

3.3.1. The Rewarding Scale

As shown in
Table 3, each of the six Rewarding items are positively correlated with each other, with coefficients ranging between .21 and .69. The strongest relationship is between item two and six, with the weakest between item three and item five. The mean and standard deviations are consistent across each of the six items.

Table 3: Intercorrelations, Descriptive Statistics & Internal Consistency Measures of the Rewarding Items

Item	R1	R2	R3	R4	R5	*M*	*SD*	α
R1	—					2.88	.77	.74
R2	.36	—				2.64	.89	.72
R3	.44	.34	—			2.69	.85	.76
R4	.49	.43	.42	—		2.54	.82	.73
R5	.29	.24	.21	.37	—	2.14	.74	.78
R6	.31	.69	.29	.38	.27	2.79	.93	.74

Note: N = 2,318; *M* = Mean; *SD* = Standard Deviation; α = Cronbach's alpha if the item is removed from the scale. All correlations are statistically significant at $p < .001$ (two-tailed) level.

3.3.2. The Willing Scale

As shown in **Table 4**, the six Willing items are positively correlated (coefficients range between .15 & .50), except for item two and item five which are uncorrelated. As with the Rewarding items, there is little variation in the mean score and standard deviation across the Willing items.

Table 4: Intercorrelations, Descriptive Statistics & Internal Consistency Measures of the Willing Items

Item	W1	W2(R)	W3	W4	W5	*M*	*SD*	α
W1	—					3.22	.65	.64
W2 (R)	.30	—				3.00	.72	.70
W3	.31	.19	—			2.79	.76	.62
W4	.34	.24	.50	—		2.91	.83	.62
W5	.15	-.03*	.33	.26	—	2.66	.83	.70
W7	.49	.27	.30	.33	.19	2.98	.78	.63

Note: N = 2,312. (R) = Reverse scored item; *M* = Mean; *SD* = Standard Deviation; α = Cronbach's alpha if the item is removed from the scale. All correlations are statistically significant at $p < .001$ (two-tailed) level, those marked * ($p > .16$, two-tailed).

3.3.3. The Able Game

Able uses MindX's gamified assessment of cognitive ability. This assessment contains two games: *Shape Dance* (an abstract reasoning task) and *Disco Numbers* (a numerical reasoning & working memory task). This assessment of cognitive ability provides an engaging and interactive element to the RAW assessment.

The objective of both games is to provide correct solutions to the problems. Each game contains 30 levels and becomes more difficult as participants progress through the game. Participants begin each game at level 10, with a "two up one down" algorithm determining how participants move from one level to the next. For example, if a participant starts on level 10 and provides a correct answer, they will progress to level 12. If they then provide an incorrect answer, they will go down to level 11, whereby they are presented with an easier problem to solve.

In both games, performance is measured across four scores: the number of rounds completed in the game session (*Rounds Completed)*, the number of correct responses made by a participant (*Correct Responses)*, the highest level reached by a participant (*Max Level*), and the proportion of correct responses to incorrect responses (*Win Proportion*). In the case of the RAW assessment, we computed the mean for each of the four performance scores.

3.3.3.1. Descriptive Statistics

Table 5 contains descriptive statistics regarding the scores associated with the two MindX games. On average, participants completed 9.74 rounds, made 4.70 correct responses,

reached a maximum level of 13.74, and an average proportion of wins to losses of .49. When creating these composite variables, two met acceptable levels of internal consistency (*Average Correct Response* & *Average Max Level*). All variables except *Average Win Proportion* were found to have issues regarding skewness and Kurtosis. For example, *Average Max Level* has a skewness statistic less than -1.00 and a leptokurtic distribution. *Average Rounds Complete* and *Average Correct Response* also had a leptokurtic distribution.

Table 5: Descriptive Statistics for the MindX Assessment

Performance Scores	*Min*	*Max*	*M*	*SD*	Skewness	Kurtosis	α
Average Rounds Completed	.50	21.50	9.74	1.89	.58	5.57	.42
Average Correct Response	.00	7.00	4.70	1.19	-.93	1.27	.63
Average Max Level	.00	19.50	13.74	3.12	-1.31	2.38	.64
Average Win Proportion	.00	.80	.49	.13	-.67	.70	.46

Note: N = 498; Min = Minimum score; Max = Maximum score; *M* = Mean; *SD* = Standard Deviation; α = Cronbach's alpha.

3.3.3.2. Construct Validity of The Gamified Assessment

To identify the MindX score that is the most valid measure of cognitive ability, we correlated participant's scores against their performance on Raven's Progressive Matrices (RPM;N = 268; Raven, Raven, & Court (1998); for more information about the RPM see **Chapter 4.3**). **Table 6** contains the results of these analyses. The majority of the MindX scores were positively correlated with the RPM. Notably, the Average Win Proportion score yielded the highest correlation with RPM scores (r = .42). As such, this score was selected for further investigation due to its low internal consistency (α = .46).

Table 6: Correlations between the Gamified Assessment & Raven's Progressive Matrices

MindX	Raven's Progressive Matrices
Average Rounds	-.13*
Average Correct	.33
Average Max Level	.35
Average Win Proportion	.42

Note: N = 268. All correlations are statistically significant at $p < .001$ (two-tailed) level, those marked * ($p > .05$, two-tailed).

To explore the suitability of a composite win score, we conducted a parallel analysis, which suggested that one component could be extracted from the data. Accordingly, we then conducted a principal components analysis, which produced a single component that explained 70% of the variance, with both win proportion variables having a factor loading of greater than .80. Accordingly, the Average Win Proportion score was deemed to be an appropriate measure of cognitive ability, and therefore a suitable Able measure.

3.3.3.3. Additional Validity Investigations

Using the Average Win Proportion score as the Able measure, we conducted additional research to investigate its relationship with other measures of cognitive ability. Specifically, participants completed *The Hogan Business Reasoning Inventory* (HBRI; Hogan, Barrett, &

Hogan, 2009) and the International Cognitive Ability Resource-Short Form (ICAR; Condon & Revelle, 2014). The HBRI is a 24-item contextualized measure of tactical and strategic thinking, while the ICAR is a 16-item inventory that measures four cognitive abilities: verbal reasoning, letter and numerical reasoning, matrix reasoning, and mental rotation.

A sample of 162 participants were recruited using Amazon's MTurk platform. The sample ranged between 20 and 64 years old (*M* = 34.88, *SD* = 10.18), with 68 participants being male and 94 being female. All were from the United States, with the majority being White (83.34%), working in non-management positions (66%), and having completed an undergraduate education (43.20%). Participants were compensated for their time and were fully debrief upon completion.

Table 7 contains the correlations between the MindX assessment and HBRI scores, and **Table 8** contains the correlations between the MindX assessment and ICAR scores. In both instances, Average Win Proportion scores had the strongest correlations with the measures of cognitive ability. Specifically, the total HBRI score (*r* = .50) and the total ICAR score (*r* = .57). These findings further support our decision to use the Average Win Proportion score as a measure of cognitive ability.

Table 7: Correlations between the Gamified Assessment & The Hogan Business Reasoning Inventory

MindX	HBRI – Total	HBRI – Tactical	HBRI - Strategic
Average Rounds	-.12*	-.09*	-.11*
Average Correct	.44	.36	.38
Average Max Level	.46	.35	.42
Average Win Proportion	.50	.41	.44

Note: N = 162. All correlations are statistically significant at $p < .001$ (two-tailed) level, those marked * ($p > .05$, two-tailed).

Table 8: Correlations between the Gamified Assessment & The International Cognitive Ability Resource

MindX	Total	Verbal Reasoning	Letter & Number Reasoning	Matrix Reasoning	Mental Rotation
Average Rounds	-.10*	-.12*	-.07*	-.13*	.04*
Average Correct	.51	.39	.46	.34	.32
Average Max Level	.53	.40	.48	.35	.34
Average Win Proportion	.57	.46	.52	.40	.30

Note: N = 162. All correlations are statistically significant at $p < .001$ (two-tailed) level, those marked * ($p > .05$, two-tailed).

4. Validity

4.1. Construct Validity

Validity concerns the degree to which scores predict meaningful non-test behavioral outcomes. That is, the validity of the RAW assessment depends on finding relationships between scores on each component of the model and data from other sources, including other assessments.

This chapter describes the construct validity of the RAW dimensions by comparing them with three different assessments.

4.2. Procedure & Sample

We described the sample used to explore the construct validity of the RAW assessment in **Chapter 3.1**. Hosted via an online platform, participants provided information regarding their demographics, then completed the RAW assessment and three other psychological assessments. Participants received compensation for their time, and all participants completed the assessments as part of low-stakes testing where results did not impact hiring, promotion, or any other personnel decisions.

4.3. Instruments

HPI. The Hogan Personality Inventory (Hogan & Hogan, 2007) contains seven scales: Adjustment, Ambition, Sociability, Interpersonal Sensitivity, Prudence, Learning Approach and Inquisitive. These scales concern the "bright side" of personality, that is, how people behave when they are at their best. There is considerable peer-reviewed research demonstrating the reliability and validity of the HPI. Please refer to its technical manual for more information (Hogan & Hogan, 2007).

Big Five Mini-Markers. Saucier's (1994) inventory contains 40 adjectives covering the Big Five taxonomy. Participants state the extent to which the marker adjectives describe them. Like the HPI, this inventory seeks to assess an individual's typical personality. We administered the inventory to a subset of our sample, with concurrent data being only available for the Rewarding (N = 157) and Willing (N = 138) scales.

Vocational Interests Inventory. Based on Holland's (1966) theory of vocational interests, this assessment consists of six scales: Realistic, Investigative, Artistic, Social, Enterprising and Conventional. Each scale contains five items and evaluates measures the extent to which individuals fit one of these vocational categories. This short inventory is based on the Hogan Motivations, Values and Preferences Inventory (MVPI; Hogan & Hogan, 1996). In the appendix, **Table 23** shows that the six scales have acceptable psychometric properties. Specifically, scores are normally distributed and have acceptable levels of internal consistency. Meta-analytic research shows that vocational interests significantly predict job performance and employee turnover (Van Iddekinge, et al., 2011).

RPM. Raven's Advanced Progressive Matrices is a 60-item assessment of fluid intelligence. It requires participants to solve geometric patterns by identifying the correct solution from a list of alternatives. The assessment is widely used in employee selection situations (Raven, et al., 1998)

4.4. The Construct Validity of the RAW Assessment

As a first step in investigating the validity of the RAW assessment, we examined correlations between the three RAW dimensions. These correlations are contained in **Table 9**.

The Rewarding and Willing scales are slightly but positively correlated (r = .14). This suggests that if people are confident, relaxed and emotionally stable (Rewarding), they are only slightly likely to be motivated and organized (Willing). Furthermore, Able is was unrelated to the Rewarding and Willing scales. These findings suggest that a total employability score should not be computed, instead researchers and practitioners should treat each dimension of the RAW assessment separately.

Table 9: Intercorrelations between the Hogan RAW Assessment

	Rewarding	Willing
Rewarding	—	
Willing	.14	—
Able	-.07*	.01*

Note: All correlations are statistically significant at $p < .001$ (two-tailed) level, except those marked * ($p > .05$, two-tailed).

Table 10 contains the correlations between the RAW assessment and the HPI — an omnibus measure of "normal" personality. The Rewarding scale is most correlated with Adjustment and Ambition, and to a lesser extent Interpersonal Sensitivity and Sociability. These findings are expected, as we define Rewarding as the tendency to be emotionally stable, confident, friendly and considerate.

The Willing scale is correlated with Prudence, Learning Approach and Ambition, while unrelated to Sociability. Given that Willing scale is designed to predict the degree to which people seem organized, committed, and eager to do and learn more, these correlations are expected.

Able scores are unrelated to all HPI scales, except Inquisitive, with which it is weakly correlated. This shows that the gamified Able dimension is independent of personality.

Table 10: Correlations between the Hogan RAW Assessment & the Hogan Personality Inventory

	Rewarding	Willing	Able
Adjustment	.68	.18	-.05*
Ambition	.56	.21	-.06*
Sociability	.31	-.01*	-.08*
Interpersonal Sensitivity	.32	.18	.00*
Prudence	.21	.28	-.03*
Inquisitive	.20	.13	.14
Learning Approach	.13	.26	.03*

Note: All correlations are statistically significant at $p < .001$ (two-tailed) level, except those marked * ($p > .05$, two-tailed).

Table 11 contains the correlations between the Rewarding and Willing scales, and the Big Five mini-markers (Saucier, 1994). For simplicity, only correlations greater than .30 are presented. Individuals with high Rewarding scores seem emotionally stable, calm and relaxed, but not depressed, anxious or moody. Individuals with high Willing scores seem organized, efficient and hard-working, but not sloppy, careless and ineffective. Together, these correlates provider further evidence for the construct validity of the Rewarding and Willing scales.

Table 11: Mini-Marker Correlates for the Rewarding & Willing Scales

Rewarding		**Willing**	
Emotionally stable/not easily upset	.44	Organized	.47
Remains calm in tense situations	.37	Does things efficiently	.41
Relaxed	.37	Efficient	.37
Relaxed/handles stress well	.36	Makes plans and follows through	.35
		Does a thorough job	.31
Is depressed/blue	-.45	Disorganized	-.45
Worries a lot	-.41	Tends to be disorganized	-.45
Moody	-.41	Inefficient	-.43
Can be moody	-.40	Sloppy	-.33
Can be tense	-.38	Can be somewhat careless	-.32
Gets nervous easily	-.36		
Temperamental	-.32		
Fretful	-.31		

Note: Rewarding N = 157. Willing N = 138. All correlations are statistically significant at $p < .01$ (two-tailed).

Table 12 contains correlations between the RAW dimensions and Holland's six vocational interests. Individuals with high Rewarding and Willing scores are most likely to be interested in Conventional, Realistic and Enterprising vocations. Able scores were unrelated to vocational interests, except negatively, albeit weakly, with Artistic and Enterprising interests.

Table 12: Correlations between the Hogan RAW Assessment & Vocational Interests

	Rewarding	Willing	Able
Realistic	.20	.12	.04*
Investigative	.17	.21	.07*
Artistic	.03*	.11	-.13
Social	.05	.16	-.05*
Enterprising	.19	.26	-.14
Conventional	.23	.49	-.02*

Note: All correlations are statistically significant at $p < .001$ (two-tailed) level, except those marked * ($p > .05$, two-tailed).

Finally, **Table 13** shows that only Able is correlated with RPM scores.

Table 13: Correlations between the Hogan RAW Assessment & Raven's Progressive Matrices

RAW	Raven's Progressive Matrices
Rewarding	.01*
Willing	.02*
Able	.42

Note: All correlations are statistically significant at $p < .001$ (two-tailed) level, except those marked * ($p > .05$, two-tailed).

Overall, the analyses outlined in this chapter support the view that the RAW assessment is a valid measure of individual differences in being Rewarding, Willing, and Able.

5. Analysis of Adverse Impact for the Hogan RAW Assessment

For organizations interested in the Hogan RAW assessment, it is important to know about potential Adverse Impact (AI). This helps ensure that decisions based on our assessment do not disproportionately impact employees because of age, sex, or race/ethnicity. This chapter defines AI and presents results for simulations using our model to identify high-potential employees. For a more detailed description of this topic, various methods for examining AI, case law from relevant court decisions, and group differences on the Hogan assessments, please consult our Adverse Impact White Paper or request a copy from your Hogan consultant.

5.1. Defining Adverse Impact

The *Uniform Guidelines on Employee Selection Procedures* (Equal Employment Opportunity Commission, Civil Service Commission, U.S. Department of Labor, 1978) defines AI as "a substantially different rate of selection in hiring, promotion, or other employment decisions which works to the disadvantage of members of a race, sex or ethnic group" (see section 1607.16). In examining the potential for AI, the *Uniform Guidelines* outlines the four-fifths rule, stating that the "selection rate for any race, sex or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact." (1978, see section 1607.4 D). Because the Age Discrimination in Employment Act (ADEA) of 1967 prohibits discrimination in selection contexts against individuals 40 years of age or older, courts have also applied this rule to cases involving potential age discrimination.

Organizations are not required to conduct validity studies for selection procedures where no AI exists. However, best practices require examining the potential for AI and accumulating validity evidence for each step of any selection process. In such efforts, statistical significance tests for mean group differences on assessment scale scores is often informative, but does not provide evidence of AI when a selection profile includes multiple assessment scales. As such, organizations must examine AI at the point at which selection decisions are made rather than differences on individual assessment scales underlying a competency or dimension.

5.2. Adverse Impact and the Hogan RAW Assessment

To investigate the potential for AI in our assessment, we first examine mean group differences based on age, sex, and race/ethnicity across the three RAW dimensions. After this, we present the results of our AI analyses testing the four-fifths rule.

5.2.1. Age Differences

Table 14 shows mean differences for the three dimensions based on age groups. Because the Age Discrimination in Employment Act (ADEA) targets individuals 40 years of age or older as needing protection, we compare scores of participants under 40 to those 40 and above. We used *t*-tests to examine mean score differences, and Cohen's *d* effect sizes (Cohen, 1988) to evaluate the practical meaning of those differences.

Table 14: Mean Differences on the Hogan RAW Assessment by Age

Dimensions	Under 40		Over 40			
	M	SD	M	SD	*t*	*d*
Rewarding	15.47	3.51	16.16	3.28	4.40***	.20
Willing	17.54	2.89	17.68	2.87	1.00	.05
Able	.50	.12	.46	.12	-3.28***	.33

Note: M = Mean; *SD* = Standard Deviation; * *t* statistic is significant at .05 level; ** *t* statistic is significant at .001 level; *** t statistic is significant at .0001 level; *d* = Cohen's *d* effect size (.00 - .19 = negligible; .20 - .49 = small; .50 - .79 = moderate; .80< = large).

There were significant age differences in Rewarding and Able scores. Specifically, individuals over 40 had significantly higher Rewarding scores, whereas individuals under 40 had significantly higher Able scores. In both cases, these differences were not practically meaningful: Cohen's *d* ranged between .20 and 33, indicating only small differences.

5.2.2. Sex Differences

Table 15 contains mean scores based on participants' reported sex. Consistent with existing legal and professional guidelines, we treat females as the protected group. We used *t*-tests to examine mean score differences, and Cohen's *d* effect sizes to evaluate their practical meaning.

Table 15: Mean Differences on the Hogan RAW Assessment by Sex

Dimensions	Male		Female			
	M	SD	M	SD	*t*	*d*
Rewarding	16.42	3.25	14.98	3.50	-9.91***	.42
Willing	17.19	2.86	17.94	2.87	6.05***	.26
Able	.51	.13	.47	.12	-3.66***	32

Note: M = Mean; *SD* = Standard Deviation; * *t* statistic is significant at .05 level; ** *t* statistic is significant at .001 level; *** t statistic is significant at .0001 level; *d* = Cohen's *d* effect size (.00 - .19 = negligible; .20 - .49 = small; .50 - .79 = moderate; .80< = large).

Across the three dimensions, we found statistically significant sex differences. Males had higher Rewarding and Able scores, whereas females had higher Willing scores. The practical implications for these differences is small: Cohen's *d* is less than .42.

5.2.3. Race/Ethnicity Differences

Table 16 provides mean group differences based on race/ethnicity. The EEOC compliance manual (Office of Management and Budget, 2006) defines the following race/ethnicity groups: (a) American Indian or Alaska Native, (b) Asian, (c) Black or African-American, (d) Hispanic, (e) Native Hawaiian or Other Pacific Islander, and (f) White. Consistent with these guidelines, we compare scores of White participants as the majority group to participants from all other racial/ethnic categories as minority groups. Because our sample provided sufficient data for White, Black/African-American, Asian participants and Hispanic individuals, we report results for those groups only. We used ANOVAs and post-hoc t-tests tests to examine the statistical significance of mean score difference, and Cohen's *d* effect sizes to evaluate the practical meaning of those differences.

Across the RAW assessment, racial/ethnic differences were found for the Rewarding and the Able dimensions. There were significant differences in Rewarding scores between White, Asian and African American participants; specifically, the latter two demographics had higher scores. Furthermore, White participants received significantly higher Able scores than African American participants. Referencing Cohen's *d*, the practical implications when considering such differences in Rewarding scores was small, while the practical differences in Able scores were moderate.

Table 16: Mean Differences on the Hogan RAW Assessment by Race/Ethnicity

Dimensions	White		*F*	Asian				African American				Hispanic			
	M	SD		M	SD	MD	*d*	M	SD	MD	*d*	M	SD	MD	*d*
Rewarding	15.50	3.44	9.63***	16.30	3.39	-.80*	.24	16.91	3.59	-1.41***	.40	16.12	3.21	-.62	.19
Willing	17.58	2.89	1.55	17.19	2.90	.39	.15	17.67	2.94	-.09	.03	18.00	2.72	-.42	.15
Able	.51	.11	8.10***	.49	.12	.02	.17	.44	.14	.07***	.57	.50	.12	.01	.08

Note: M = Mean; *SD* = Standard Deviation; * *t* statistic is significant at .05 level; ** *t* statistic is significant at .001 level; *** t statistic is significant at .0001 level; *d* = Cohen's *d* effect size (.00 - .19 = negligible; .20 - .49 = small; .50 - .79 = moderate; .80< = large).

5.2.4. Adverse Impact Analysis in the Operational Use of the Hogan RAW Assessment

The examination of mean group scores shows small or moderate differences. However, mean differences per se do not indicate adverse impact (AI). To examine the potential for AI in operational use, we need to evaluate the cutoff scores organizations might use to screen employees, based on their RAW scores. This section outlines the results of our AI analyses and we provide cutoff scores that can be used to assist selection decisions without violating the four-fifths rule.

Table 17 contains our recommended cutoff scores for the RAW assessment. Cutoff scores vary across the three dimensions to maximize the trade-off between selecting out the lowest scoring individuals while not creating adverse impact. Organizations using different cutoff scores should evaluate the potential for AI based on the specific scores they use for making or influencing hiring decisions.

Using the listed cutoff scores, we conducted AI tests for each of the three dimensions across age, sex and race/ethnicity. **Table 19** contains the results of the AI analysis for the Rewarding scale,

Table 20 for the Willing scale, and **Table 21** for the Able measure. Using the selection decision rules outlined in **Table 17**, the four-fifths rule is not violated. Accordingly, there is no AI in the RAW assessment.

Table 17: Recommended Selection Decisions Rules for Using the Hogan RAW Assessment.

	Fails to Meet Cutoff Scores		Meets Cutoff Scores	
	Raw Score	Percentile	Raw Score	Percentile
Rewarding	≤ 13	≤ 25%	≥ 13	≥ 25%
Willing	≤ 16	≤ 25%	≥ 16	≥ 25%
Able	≤ .37	≤ 15%	≥ .37	≥ 15%

Finally, we explored non-AI producing scoring profiles using Rewarding, Able, and Willing scores simultaneously. The cutoff scores in **Table 17** are on a scale-by-scale basis, however practitioners may want to make selection decisions by taking all three dimensions into account. The results of these investigations are reported in **Table 18** and in **Table 22**. Given the low cutoff scores, we recommend that selection decisions are best made on a scale-by-scale basis.

Table 18: Recommended Selection Decisions Rules for Using the Three Dimensions Simultaneously

	Fails to Meet Cutoff Scores		Meets Cutoff Scores	
	Raw Score	Percentile	Raw Score	Percentile
Rewarding	≤ 11	≤ 5%	≥ 11	≥ 5%
Willing	≤ 13	≤ 5%	≥ 13	≥ 5%
Able	≤ .27	≤ 5%	≥ .27	≥ 5%

Table 19: Selection and Adverse Impact Ratios for Using the Rewarding Scale

		Total Sample	Select-In	Selection Rate	Selection Ratio	
Sex					Male : Female	Female : Male
	Male	1,034	920	.89	1.16	.86
	Female	1,121	859	.77	NA	NA
					Under 40 : Over 40	Over 40: Under 40
Age	Under 40	1,517	1,220	.80	.92	1.09
	Over 40	638	559	.87	NA	NA
					White : Minority	Minority : White
Ethnicity	White	1,771	1,439	.81	NA	NA
	Asian	135	118	.87	.93	1.08
	African American	141	129	.91	.89	1.13
	Hispanic	108	93	.86	.94	1.06

Table 20: Selection and Adverse Impact Ratios for Using the Willing Scale

		Total Sample	Select-In	Selection Rate	Selection Ratio	
Sex					Male : Female	Female : Male
	Male	1,021	758	.74	.90	1.11
	Female	1,124	923	.82	NA	NA
Age					Under 40 : Over 40	Over 40: Under 40
	Under 40	1,502	1,161	.77	.96	1.05
	Over 40	643	520	.80	NA	NA
Ethnicity					White : Minority	Minority : White
	White	1,774	1,389	.78	NA	NA
	Asian	134	101	.75	1.04	.96
	African American	137	109	.80	.98	1.03
	Hispanic	100	82	.82	.95	1.05

Table 21: Selection and Adverse Impact Ratios for Using the Able Game

		Total Sample	Select-In	Selection Rate	Selection Ratio	
Sex					Male : Female	Female : Male
	Male	252	188	.75	1.07	.93
	Female	229	161	.70	NA	NA
					Under 40 : Over 40	Over 40: Under 40
Age	Under 40	335	308	.91	1.13	.88
	Over 40	126	101	.80	NA	NA
					White : Minority	Minority : White
Ethnicity	White	208	188	.90	NA	NA
	Asian	99	85	.86	1.04	.96
	African American	98	71	.73	1.23	.81
	Hispanic	76	65	.86	1.04	.96

Table 22: Selection and Adverse Impact Ratios for Using the Three Dimensions Simultaneously

		Total Sample	Select-In	Selection Rate	Selection Ratio	
Sex					Male : Female	Female : Male
	Male	233	206	.88	.99	1.01
	Female	210	186	.89	NA	NA
					Under 40 : Over 40	Over 40: Under 40
Age	Under 40	328	290	.88	1.00	1.00
	Over 40	115	101	.88	NA	NA
					White : Minority	Minority : White
Ethnicity	White	193	178	.92	NA	NA
	Asian	91	80	.88	1.05	.96
	African American	92	72	.78	1.18	.85
	Hispanic	67	61	.91	1.01	.99

6. References

Baruch, Y., & Bozionelos, N. (2011). Career issues. In *APA handbook of industrial and organizational psychology, Vol 2: Selecting and developing members for the organization* (pp. 67–113).

Blickle, G., Fröhlich, J. K., Ehlert, S., Pirner, K., Dietl, E., Hanes, T. J., & Ferris, G. R. (2011). Socioanalytic theory and work behavior : Roles of work values and political skill in job performance and promotability assessment. *Journal of Vocational Behavior*, *78*(1), 136–148.

Brand, J. E. (2015). The Far-Reaching Impact of Job Loss and Unemployment. *Annual Review of Sociology*, *41*, 359–375.

Cappelli, P. H. (2015). Skill gaps, skill shortages, and skill mismatches: Evidence and arguments for the United States. *Industrial and Labor Relations Review*, *68*(2), 251–290.

Cappelli, P., & Keller, J. (2014). Talent Management: Conceptual Approaches and Practical Challenges. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 305–331.

Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., Hogan, R. (2013). New Talent Signals: Shiny New Objects or a Brave New World. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *53*(9), 1689–1699.

Chamorro-Premuzic, T., & Furnham, A. (2010). *The Psychology of Personnel Selection*. Cambridge University Press.

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

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64.

De Stefano, V. (2015). The rise of the " just-in-time workforce": on-demand work, crowd work and labour protection in the "gig-economy." *Comparative Labor Law & Policy Journal*, *Forthcomin*(November 2015), 8–10.

Edwards, J. R., & Cable, D. M. (2009). The value of value congruence. *The Journal of Applied Psychology*, *94*(3), 654–77.

Ehrhart, K. H., & Makransky, G. (2007). Testing Vocational Interests and Personality as Predictors of Person-Vocation and Person-Job Fit. *Journal of Career Assessment*, *15*(2), 206–226.

Equal Employment Opportunity Commission, Civil Service Commission, U.S. Department of Labor, & U. S. D. of J. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, *43*, 38290–38309.

Ford, M. (Martin R. . (n.d.). *Rise of the robots : technology and the threat of a jobless future*. Basic Books.

Fugate, M., Kinicki, A. J., & Ashforth, B. E. (2004). Employability: A psycho-social construct, its dimensions, and applications. *Journal of Vocational Behavior, 65*(1), 14-38.

Fuller, S., & Stecy-Hildebrandt, N. (2014). Career pathways for temporary workers : Exploring heterogeneous mobility dynamics with sequence analysis.

Hall, D. T., & Chandler, D. E. (2004). Introduction : The Value of a Dual Perspective. *Journal of Organizational Behavior J . Organiz . Behav*, *25*, 1–22.

Hogan, R. T., Chamorro-Premuzic, T., & Kaiser, R. B. (2013). Employability and Career

Success : Bridging the Gap Between Theory and. *Industrial and Organizational Psychology*, *6*, 3–16.
Hogan, R., Barrett, P., & Hogan, J. (2009). *Hogan Business Reasoning Inventory Manual*. Tulsa, OK: Hogan Assessment Systems.
Hogan, R., & Hogan, J. (2007). *Hogan Personality Inventory Manual* (3rd ed.). Hogan Assessment Systems.
Hogan, R., & Hogan, J. (2009). *Hogan Development Survey manual* (2nd ed.). Tulsa, OK.
Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. *Psychological Assessment Resources.*
Holland, L., J. (1966). The psychology of vocational choice: A theory of personality types and model environments. Oxford, UK: Blaisdell.
Huws, U., & Joyce, S. (2016). Crowd Working Survey: Size of the UK's "Gig Economy" revealed for the first time. *Foundation for European Progressive Studies*.
Judge, T. A., Higgins, C. A., 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.
Koenig, J. A. (2011). *Assessing 21st Century Skills: Summary of a Workshop*. *National Academies Press*.
Kuncel, N. R., Ones, D. S., & Sackett, P. R. (2010). Individual differences as predictors of work, educational, and broad life outcomes. *Personality and Individual Differences*, *49*(4), 331–336.
Lashinsky, A. (2012). Why employment matters. Retrieved from http://fortune.com/2012/10/12/why-employment-matters/
Levine, L. (2013). The increase in unemployment since 2007: Is it cyclical or structural?. *Congressional Research Service*, *7–5700*(R41785), 1–17.
Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, *97*(2), 460–468.
Lindqvist, E., & Westman, R. (2010). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics, 3*(1), 101–128.
Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of Cognitive Assessment and Cognitive Training: A Systematic Review of Applications and Efficacy. *JMIR Serious Games*, *4*(2), e11. https://doi.org/10.2196/games.5888
Mikulincer, M., Shaver, P. R., Simpson, J. A., & Dovidio, J. F. (2015). *APA handbook of personality and social psychology*. *APA handbook of Personality and Social Psychology, Volume 3: Interpersonal Relations.* (Vol. 3).
Milner, A., Page, A., & LaMontagne, A. D. (2013). Long-Term Unemployment and Suicide: A Systematic Review and Meta-Analysis. *PLoS ONE*, *8*(1), e51333.
Molloy, R., Smith, C. L., Trezzi, R., & Wozniak, A. (2016). Understanding Declining Fluidity in the US Labor Market. *Brookings Papers on Economic Activity*, 183–259.
Nauta, M. M. (2010). The development, evolution, and status of Holland's theory of vocational personalities: Reflections and future directions for counseling psychology. *Journal of Counseling Psychology*, *57*(1), 11–22.
Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Pedictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, *58*(2), 367–408.
Office of Management and Budget. (2006). *Section 15: Race and Color Discrimination. EEOC*

compliance manual. Washington, DC: Government Printing Office.
Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, *104*(11), 3565–3599.
Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*.
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 : A Journal of the Association for Psychological Science*, *2*(4), 313–45.
Roelfs, D. J., Shor, E., Davidson, K. W., & Schwartz, J. E. (2011). Losing life and livelihood: A systematic review and meta-analysis of unemployment and all-cause mortality. *Social Science and Medicine*, *72*(6), 840–854.
Saucier, G. (1994). Mini-Markers: A Brief Version of Goldberg's Unipolar Big-Five Markers. *Journal of Personality Assessment*, *63*(3), 506–516.
Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, *71*(3), 432–439.
Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B. W. (2016). Vocational Interests Assessed at the End of High School Predict Life Outcomes Assessed 10 Years Later Over and Above IQ and Big Five Personality Traits. *Journal of Personality and Social Psychology*.
The Secretary's Commission on Achieving Necessary Skills. (1991). What work requires of schools: A SCANS report for America 2000. *America*, *2008*(June 1991), 52.
Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology*, *96*(6), 1167–1194.
Wille, B., De Fruyt, F., Dingemanse, S. A., & Vergauwe, J. (2015). A closer look at the psychological diversity within Holland interest types: Construct validation of the Career Insight Questionnaire. *Consulting Psychology Journal: Practice and Research*, *67*(3), 234–257.

7. Appendix

Table 23: Bivariate Correlations & Classical Scale Statistics for the Vocational Interests Inventory

	1.	2.	3.	4.	5	*M*	*SD*	Skewness	Kurtosis	α
1. Realistic	—					13.11	3.12	-.21	-.22	.73
2. Investigative	.58	—				14.55	2.85	-.31	.10	.77
3. Artistic	-.08	.16	—			12.70	3.03	-.09	-.25	.70
4. Social	-.03*	.22	.33			14.09	2.67	-.17	.26	.72
5. Enterprising	.23	.20	.15	.00*	—	12.47	2.67	.09	.05	.66
6. Conventional	.34	.27	.04*	.05	.37	12.94	2.68	-.03	.23	.65

Note: Realistic N = 2334; Investigative N = 2324; Artistic N = 2332; Social N = 2347; Enterprising N = 2344; Conventional N = 2332. Scores can range between 5 and 20. All correlations are statistically significant at $p < .01$ (two-tailed) level, those marked * ($p > .05$, two-tailed).