

TECHNICAL MANUAL & USER GUIDE

WORK PROFILE

This manual has been created by the psychologists from the Research and Innovation department of Central Test.

The assessment WORK PROFILE and the Central Test logo are registered trademarks.

No reproduction, modification or adaptation is permitted without the written consent of the publisher Central Test.

*Copyright © 2020 by Central Test - Empowering Talent.
All rights reserved.*

Table of Contents

Introduction	4
1. General Description and Applicability	5
1.1. Main Areas of Application	6
1.2. Questionnaire Structure	7
1.3. Mode of Administration, Languages and Special Requirements	7
1.4. Social Desirability	8
1.5. Administration Instructions	8
1.6. Technical and Security Issues	9
1.6.1. Respecting the Ethical Charter	9
1.6.2. Confidentiality and Security of User Data	10
1.6.3. Hardware Architecture and Server Availability	10
1.6.4. Computer Generated Reports	10
1.7. Supply, Conditions and Availability	11
2. Conceptual and Practical Scale Development	12
2.1. Theoretical Background	12
2.1.1. Theories used in Work Profile	13
2.2. Development of Work Profile	18
2.2.1. Dimensions, Factors and Item Construction	18
2.2.2. Ipsative Response Format	19
2.2.3. Normative Response Format	20
2.2.4. Thurstonian Item Response Theory (IRT) and Ipsative Responses	20
2.2.5. Pretests and Pilot Study	22
3. Scoring	23
3.1. The Relationship between Raw Scores and Norms	23
3.2. Transforming Raw Scores to Standardised Scores	25
3.5. Creating In-house Norms	26
3.6. Link between Scores and Report Comments	26
4. Sample Descriptives	28
5. Norms	29
6. Reliability	38
6.1. Internal Reliability	38
6.2. Test-retest Reliability	39
7. Validity	41
7.1. Concurrent Validity	41
7.1.1. The Mini-IPIP Five-factor Personality Scale	41

7.2. Construct Validity	42
7.2.1. Inter-factor Correlations	43
7.2.2. Confirmatory Factor Analysis	45
8. WORK PROFILE and Equal Opportunities	47
8.1. Measurement Equivalence	47
8.1.1. Differential Item Functioning (DIF)	47
8.1.1.a. DIF Analysis: Gender	48
8.1.1.b. T-tests and ANOVAS	49
9. Additional Information and Considerations	53
9.1. Interpreting Scores	53
9.2. Organisational Need: Career Selection, Recruitment and Development	54
References	55

Introduction

WORK PROFILE is a psychometric evaluation tool designed to analyse workplace, social and emotional traits as well as the aspirations and values of individuals. It can be administered at all professional levels; however, it was especially conceived for lower skill level professions.

The labour market has changed dramatically over the course of the past several decades. Little more than a generation ago, lower skill level jobs, more often than not, did not require academic or social competencies. Today however, that premise may not be so easily defended. The job opportunities that require little proficiency are rapidly changing as there is a shift from a 'production driven' economy to a 'service driven' economy. This shift now requires that prospective employees be more adept than those of prior generations with regards to 'new' workplace demands (Thompson, 2001). Thus, employers now deem it essential to consider the personality traits, motivations and values as well as the skills of the employees they recruit.

In light of this, **WORK PROFILE** was constructed with the objective of creating a straightforward and easily comprehensible personality assessment, to be accessible to companies for the employment process of all qualification level jobs, including lower level. **WORK PROFILE** is based on personality research, such as Costa & McCrae's "OCEAN Model" (McCrae, R. R., & John, O. P., 1992), often referred to as the "Big 5 Personality Traits." This model states that five broad domains constitute an individual's personality. **WORK PROFILE** comprises of 15 personality factors, all of which can be associated with the Big 5 Model.

WORK PROFILE serves a dual purpose:

- I. For organisations, this assessment provides a quick and comprehensive profile of their applicants.
- II. For individuals, it marks the beginning of their journey towards better understanding themselves.

Some benefits of **WORK PROFILE** that most distinguish it from other personality tools in the market are the following:

- I. More contemporary and business relevant items
- II. Short average completion time of 10-12 minutes
- III. Ipsative and normative scaling items to guard against social desirability and impression management
- IV. Developed using state-of-the-art psychometric techniques such as Thurstonian Item Response Theory
- V. User-friendly and customizable reports that can be easily tailored by the administrator
- VI. Practical and easy-to-understand assessment and report

This Technical Manual and User Guide will enable you to understand the theoretical foundations on which **WORK PROFILE** is based, the 15 factors that comprise it, and the manner in which the questionnaire is designed and structured. This manual and guide will also help you familiarise yourself with the stages of test administration, report interpretation and applications in various organisational contexts.

1. General Description and Applicability

WORK PROFILE assesses 3 dimensions of personality - *Workplace Traits*, *Social and Emotional Traits*, and *Aspirations and Values* - which in turn cover 15 factors. These factors are considered to be fundamental in providing insight into an individual's personality profile and professional aspirations, and in finding the right fit between a candidate's personality and a job.

A description of the dimensions along with their factors is provided below.

Workplace Traits

This dimension encompasses those factors which fall under workplace traits. Consisting of *Organisation*, *Compliance*, *Dynamism*, *Sociability* and *Sense of Responsibility*, these are factors that can be characterised as behaviours which commonly predict performance in the workplace. A high score on a workplace trait represents an individual's natural tendency to engage in the behaviour as defined by the factor.

1. **Organisation** refers to being methodical at work and planning one's activities in an orderly manner. An individual with such a tendency is likely to keep the environment around them structured, and ensure that the methods and steps involved in their tasks are clearly laid out and precise.
2. **Compliance** refers to respecting and following established guidelines and standards set by the organisation. An individual with such a tendency will respect the reason behind why procedures and rules are put into place, and will therefore work according to them.
3. **Dynamism** refers to an individual's tendency to be proactive, and achieve fast results with high energy. Such individuals are likely to work vigorously, and will want to progress relatively quickly from one task to another.
4. **Sociability** refers to an individual's tendency to socialise and get to know those around them at work. Such individuals will take the time to form bonds and create personal relationships with those that they work with, and will feel comfortable doing so.
5. **Sense of Responsibility** refers to an individual's tendency to feel that it is their responsibility to take on more initiatives at work. Such individuals are likely to take on more tasks than what is expected of them, by seeking work that is additional to their job role.

Social and Emotional Traits

This dimension refers to those traits which are essential to an individual's understanding and management of their social and emotional behaviours. This encompasses the factors *Persuasiveness*, *Self-Control*, *Tenacity*, *Assertiveness* and *Adaptability*. A high score on a social and emotional trait represents an individual's natural tendency to engage in the behaviour as defined by the factor.

1. **Persuasiveness** refers to an individual's tendency to adapt their speech in order to tactfully convince the person they are talking to. Such individuals are likely to present their ideas in a way that their audience will find attractive, in order to win them over.
2. **Self-Control** refers to an individual's tendency to compose their emotions during stressful or emotionally difficult situations. Such individuals will be in control of their feelings, and are

able to compose themselves without reacting impulsively, should they experience negative reactions such as frustration.

3. **Tenacity** refers to an individual's perseverance to continue working towards their objectives, even when faced with difficulties. An individual with such a tendency will maintain their determination, should obstacles arise in their work, and not get discouraged easily by failure.
4. **Assertiveness** refers to an individual's tendency to articulate their opinions, even when faced with disagreement. Such individuals will express their viewpoints, and ensure that they are heard, despite the possibility of being opposed by conflicting ideas.
5. **Adaptability** refers to an individual's tendency to change their habits easily, and to be able to adjust to altering environments, by breaking routines and finding novel solutions when faced with uncertainty.

Aspirations and Values

This dimension relates to the aspirations and values that an individual holds, which drive their occupational choices and behaviours. Consisting of *Sense of Service*, *Ambition*, *Autonomy*, *Intellectual Curiosity* and *Leadership*, these factors are characterised by an individual's preference towards the behaviour defined by each factor. A high score on an aspiration and value indicates that the person enjoys engaging in the related factor.

1. **Sense of Service** refers to a preference for engaging in activities that are helpful to other people. An individual with such an orientation will want to dedicate their time to others and will enjoy being able to attend to the well-being of those in need.
2. **Ambition** refers to the strong desire for success and working towards greater and bigger heights. An individual with such a preference will want to keep reaching high goals, and will enjoy involving themselves in activities that permit them to further themselves and achieve more targets.
3. **Autonomy** refers to the preference of being independent at work. An individual who endorses this factor will want to make their own decisions at work, choose their own working methods and work with little supervision in order to have a say over how they conduct their work.
4. **Intellectual Curiosity** refers to an individual's desire for learning new topics and educating themselves on concepts that they do not have an understanding of. Such individuals enjoy being able to take part in tasks that enable them to acquire new knowledge and broaden their intellectual horizons.
5. **Leadership** refers to an individual's preference for being directive. Such individuals will enjoy instructing others and delegating tasks at work, and are likely to feel comfortable in authoritative positions.

1.1. Main Areas of Application

For recruitment and career development opportunities, the criteria for candidate selection are diverse. Personality and motivations nowadays are as important as skills in predicting performance. They are crucial to take into account for employee selection, promotion and development.

WORK PROFILE measures the aforementioned factors in order to give human resource professionals a global perception of a candidate or employee's profile. It has also been designed to relate specifically to the employment process of lower skill level jobs.

The questionnaire is quick, simple and easily understood by all skill-levels, including those candidates that may not speak English natively. Furthermore, the report can be easily interpreted by non-psychologists.

1.2. Questionnaire Structure

WORK PROFILE consists of 60 core questions and 30 normative questions. Questions are presented in a forced-choice format, whereby each question is comprised of two statements which require the candidate to choose one option that best represents their behavioural tendencies or preferences.

Each statement assesses different components of the three dimensions measured in the assessment and each of the 15 factors is measured against each other twice, within their dimensions. The core items are paired in such a way that each response option assesses a different factor. The normative items are paired in such a way as to directly assess an individual's preference or behavioural tendencies of the factor, against the factor's opposing behaviour (corresponding to a lack of the factor).

Even though individuals tend to have more experience with Likert scale work assessments in which they can position themselves in relation to a trait, the questionnaire structure of **WORK PROFILE** facilitates both a pleasant test-taking experience as well as better face validity of results. This item structure also reduces the likelihood of respondent fatigue and response bias. Furthermore, forcing respondents to choose one of two options safeguards against impression management and the candidate's attempts to fake results. The examples below depict the questionnaire structure of **WORK PROFILE**.

An example of a Core Item:

- ☐ I enjoy learning at work.

- ☐ I enjoy succeeding in my career.

An example of a Normative Item:

- ☐ I am comfortable when things remain as they are.

- ☐ I am comfortable when my environment varies.

1.3. Mode of Administration, Languages and Special Requirements

WORK PROFILE is administered online and works with all computer systems including Windows, MAC-OS and Linux. A PDF version is available if required. The assessment can be administered remotely and without supervision. However, should standardised group assessments be required, this can be accommodated. As reports are always generated online, if the test has been administered in paper format, then the scores must manually be entered onto the online platform to generate candidate reports.

WORK PROFILE was designed to be easily used by individuals from diverse backgrounds and for those with a range of special needs. All test-takers are able to adjust visual screen specifications based on individual requirements. Similarly, respondents are able to access test instructions throughout the assessment. It is also possible to navigate through questions or change answers as desired. Alternatively, for blind candidates or those with severe visual disabilities, items may be read aloud by a trained administrator. Individuals with hearing impairments are not excluded from taking the questionnaire, since all content is presented visually.

Although the average time for completion is 10 to 12 minutes, **WORK PROFILE** has no time limit. Therefore, individuals with learning disabilities are able to take the time they need. Please contact the test publisher for guidance on administering the questionnaire to individuals or groups with special requirements.

1.4. Social Desirability

Social desirability is the tendency to present oneself in a manner that will be viewed favourably by others. When done unintentionally, it is referred to as self-deception. When done intentionally, it is referred to as impression management (Tett & Simonet, 2011). As people are social by nature, they are generally inclined to seek some degree of social acceptance. Unlike aptitude tests, the results of questionnaires like **WORK PROFILE** are affected by impression management. An individual's attempt to present himself/herself in a favourable light negatively impacts test results, whether the questionnaire is completed online or in paper-based format (Dodou & de Winter, 2014).

WORK PROFILE actively guards against social desirability with the use of an ipsative (forced choice) questionnaire format. Research has found ipsative measures to be more resistant to distorted responding compared to more commonly used Likert scales (Bowen, Martin, & Hunt, 2002; Kluger, Reilly, & Russell, 1991). Not only does the ipsative format reduce the candidate's ability to inflate their scores, but items are also combined in such a manner that makes it difficult for candidates to conceive a response strategy, as items are paired with equally appealing response options.

1.5. Administration Instructions

In most cases the assessment will be taken remotely through online access, and without supervision. However, in the case of taking the assessment within a supervised setting, the following guidelines should be applied.

Although instructions for completing **WORK PROFILE** are presented on-screen, it is important that the administrators communicate with the candidates. Prior to starting the assessment, instructions should be read aloud by the administrator, and candidates should be given the opportunity to ask questions before taking the assessment. Administrators should verify that candidates understand what they are required to do prior to starting the assessment. Test instructions are the same for every individual. Standardised instructions ensure that all participants are exposed to similar conditions during the assessment process, to ensure that it is fair and objective, and that assessment conditions do not directly or indirectly give any individual or group an advantage.

Further tips:

- If candidates have questions after the assessment session has begun, the administrator should try to answer them without disturbing other candidates.

- If a candidate does not understand a question, the administrator should answer carefully and avoid explaining the meaning of words or questions in his/her own words in order to not influence the candidate.
- If an administrator suspects that a candidate's lack of understanding is due to their education level, the assessment should be stopped.
- If an administrator suspects that a candidate's lack of understanding is due to language proficiency, the administrator should ask the candidate to take the assessment in the language they are most comfortable with.
- If an administrator suspects that a candidate is being over cautious when responding, the candidate should be encouraged to endorse their most truthful response.

Candidates receive a notification to take the questionnaire via email with an embedded web link and login details. Web links to the questionnaire are password protected and unique to the candidate's account. It is possible to set expiration dates in order to ensure that candidates complete the questionnaire within the specified time-frame. It is also possible to configure the invitation in a way that the assessment must be completed at the same IP address once the questionnaire has been accessed from a particular IP address.

WORK PROFILE has no time restriction. However, candidates generally take about 10 to 12 minutes to complete the questionnaire. For monitoring purposes, completion times are recorded to allow administrators to see if a candidate rushed through the questionnaire or needed an unusual amount of time to complete the test. Candidates are encouraged to answer questions spontaneously, but without rushing, in order to avoid spending too much time thinking about each statement as there is no right or wrong answer.

1.6. Technical and Security Issues

This chapter details guidelines governing the security of administering **WORK PROFILE**. These guidelines ensure that every administration of **WORK PROFILE** is a pleasant, objective and fair process. Security specifications ensure that candidate information is protected both during administration and after assessment completion.

1.6.1. Respecting the Ethical Charter

As a publisher of online assessment tools, *Central Test* respects the legal conventions of French Labour Law and the regulatory provisions resulting from the law of 6 January 1978 pertaining to information technology, files and freedoms: Declaration N°1834067. As of the 25th of May, 2018, *Central Test* will comply with the GDPR (General Data Protection Regulation) regulations made by the European Parliament & Council.

Test results may not be forwarded to a third person without the explicit consent of the test taker. A test taker also has the right to access their results.

The assessor must grant them access automatically or when the test taker requests it.

The test taker has the right to request that all his/her personal data be permanently deleted. To do this, he/she must contact either the organisation or *Central Test* directly, who will delete his/her data.

1.6.2. Confidentiality and Security of User Data

Central Test guarantees the security and confidentiality of data generated during the administration of online tests. Access codes are encrypted in the database and cannot be accessed by anyone. Passwords are computationally generated and are not visible to companies. A URL rewriting module (secure link generator) is installed and has transparent parameter passing for the user.

Browsing and data exchange via the *Central Test* application can be secured and represented by an icon in the form of a padlock in the internet browser. The *Central Test* application has a <<GeoTrust SSL TRUE BusinessID>> digital signature, issued and certified by GeoTrust®. This certificate is proof we take security very seriously, and protect all *Central Test* assessments from phishing attacks. Invitations to take tests are made via secure links that are of limited use (in terms of number and duration). The tests are timed and IP addresses are tracked. Access to questionnaires may also be limited in terms of number and duration.

1.6.3. Hardware Architecture and Server Availability

Central Test's architecture consists of a network of 10 physical dedicated servers.

Each server, providing high quality service, is housed in a specialised data centre featuring:

- Dedicated lines
- A guaranteed bandwidth of 100Mbps (per server)
- Redundant energy sources
- 4-hour Equipment Restoration time
- 2-hour Incident Intervention time
- Monitoring and assistance 24/7/365

Access to server administration is restricted by the IP address and is password-secured. Data exchanges are also secure.

The *Central Test* application has been tested using the following browsers:

- Internet Explorer 8 or above
- Firefox 21 or above
- Google Chrome 30 or above
- Safari 7 or above

1.6.4. Computer Generated Reports

Reports can be generated within minutes of a candidate's completion of the questionnaire. **WORK PROFILE** report is user-friendly and customisable. Report content can be tailored depending on the needs of the administrator. Separate reports may be configured for the candidate and the company, recruiter, coach or psychologist. Reports are also available in PDF format.

1.7. Supply, Conditions and Availability

Conditions of supply and availability for **WORK PROFILE** vary from country to country. However, the following points generally apply:

- **WORK PROFILE** is only available for Psychologists, Human Resource Professionals and Coaches.
- Training is compulsory prior to use and distribution.
- This manual is available to individuals who have been trained to use **WORK PROFILE**.
- **WORK PROFILE** can be ordered online, or by contacting any *Central Test* customer service representative or local partner worldwide (customer@centraltest.com).
- For information on **WORK PROFILE** prices and related products and services, contact any *Central Test* customer service representative or local partner worldwide.
- For more information on the training, please contact *Central Test*'s customer service or one of *Central Test*'s partners.

2. Conceptual and Practical Scale Development

WORK PROFILE was developed using robust theoretical concepts, text analysis of practical and relevant work requirements and advanced scientific analyses. All items in the item bank were piloted and pretested with professionals from multiple continents in order to ensure that the work behaviours assessed were contemporary, global and culture-free.

2.1. Theoretical Background

Sir Francis Galton (1884), considered to be the pioneer of psychometrics, was the first scientist to analyse the structure of personality using a lexical hypothesis. His theory proposed that personality characteristics within people's lives will become an important part of their language. Building on this hypothesis, Gordon Allport (1937) generated a list of 17,953 unique words to describe personality, leading to the coining of the term "personality" itself as well as a later clustering of terms and structuring of factors, to create personality traits by Raymond Cattell (1957) and Costa & McCrae (1985).

As the world of business became an ever-changing and increasingly important subject, psychometrics subsequently started to acquire an essential role in the process of employment. Helping to predict behaviour and performance, psychometric assessments allow recruiters access to an efficient method of narrowing down and effectively selecting possible candidates. Unlike subjective judgements which may be formed during interviews, personality and ability tests provide an objective and reliable alternative. These assessments can be designed specifically to find a desired personality trait or to match a certain working environment, and therefore to select the individual best suited to the role.

An assessment can be called psychometric as long as it is standardised, applicable across groups, and has proven reliability and validity. The link between personality and work performance has been measured thoroughly over the last 30 years, and it is long established to be both reliable and valid (Hurtz & Donovan, 2000). One study that has significantly influenced the use of personality tests in the workplace was conducted by Barrick and Mount (1991). Using the Five Factor Model (FFM), they found that personality dimensions could accurately predict at least one aspect of performance at work in a variety of jobs including professionals, police, sales staff and skilled/semi-skilled workers. Other studies later found that extraversion and conscientiousness are also related to job performance (De Fruyt & Mervielde, 1999), and that creativity is linked to higher levels of neuroticism and has a negative effect on individual work success (Rothmann & Coetzer, 2003). The use of personality assessments is crucial for any process where the individual is central to productivity (Cruise, 2012).

WORK PROFILE has been designed with the intention of providing an accurate analysis of a candidate's strong and weak areas in relation to traits required in lower skill level jobs. Getting recruitment wrong can be a high price for companies to pay. Improving the efficacy of recruitment at this level by administering a test like **WORK PROFILE** can have a significant impact. As a personality assessment, **WORK PROFILE** is a highly effective predictor of job performance, and will help recruiters select the best candidate for the role.

WORK PROFILE helps organisations to better recognise the spectrum of personality traits associated with lower skill level jobs and to pinpoint whether candidates meet their role requirements. It facilitates the recruitment process by establishing a match between job applicants and a variety of technical and trade occupations that include:

- catering assistants, building inspectors and police officers.

- machine operators, care-takers, retail, clerical and secretarial occupations.
- postal workers, hotel porters, drivers, cleaners, electricians and plumbers.

Type vs. Trait

There are two main approaches to classifying individuals in personality psychology: type and trait. Type theories categorise personality as a set of interrelated elements that make an individual unique. People are assigned to categories of personality (e.g. introversion – extraversion). Types express preferences that allow us to predict behaviour, understand changes in behaviour over time and identify how behaviours may be developed with intervention. Type theories regard opposing traits as two distinct classes and are therefore appropriate assessments for workplace development, career guidance and coaching.

Trait theories on the other hand, regard personality as degrees of expression along a continuum of behaviour. A trait could be defined as the tendency to act in a certain way and could be expressed by the behaviour of the individual in diverse situations (Nadeau, Acier, Miranda, Pervin, & John, 2004). For instance, when someone is described as a kind person it means that he/she tends to behave in a kind way, regardless of the situation. Trait approaches are commonly used in work assessments as they are effective predictors of stable behaviour. It is expected that a rigorous person will be a rigorous employee. For recruitment purposes, trait theories are recommended over type theories as they give a more detailed description of an individual's potential instead of allocating unique individuals to a limited number of personality types.

2.1.1. Theories used in Work Profile

WORK PROFILE incorporates several well-established theoretical frameworks into the development of its structure.

Five Factor Model

The Five Factor Model (FFM), which is considered to contain the core traits (Costa & McCrae, 1987) in personality measurement, plays a significant role. The FFM consists of five dimensions; *Agreeableness* - the tendency to be humanistic, cooperative and trusting of others; *Conscientiousness* - the extent to which an individual is reliable, persistent and ambitious; *Extraversion* - the inclination to be sociable, desiring to be among others, thus requiring excitement and stimulation; *Neuroticism* - the inclination to feel negative emotions including fear, guilt, sadness and anger; and *Openness to Experience* - the extent to which an individual is intellectually curious, has insight and is imaginative (John & Srivastava, 1999; Major, Turner, & Fletcher, 2006; Mooradian, Renzl, & Matzler, 2006). **WORK PROFILE** adopts this approach in measuring personality traits in the workplace.

HEXACO

Lee and Ashton (2004) proposed a 6-dimension model based on evidence of previous research on trait-based personality models. The model includes all dimensions measured by the FFM as well as the additional dimension of Honesty-Humility (H). Within the model, dimensions are classified as: *Honesty-Humility (H)*, which measures facets of sincerity, fairness, greed avoidance and modesty; *Emotionality (E)*, which assesses fearfulness, anxiety, dependence and sentimentality; *Extraversion (X)*, which identifies social self-esteem, social boldness, sociability and liveliness; *Agreeableness (A)*, which measures forgiveness, gentleness, flexibility and patience; *Conscientiousness (C)*, which assesses organisation, diligence, perfectionism and prudence; and *Openness to Experience (O)*, which examines an individual's tendency towards aesthetic appreciation, inquisitiveness, creativity and

unconventionality. For the development of **WORK PROFILE**, facets measured by the (H) dimension of HEXACO were important as an individual's values and expression of said values significantly influence role, team and organisational fit as well as his/her approach to making decisions.

FIRO

In 1958, William Schutz proposed the Fundamental Interpersonal Relations Orientation (FIRO). This states that humans have three interpersonal needs; a need for inclusion, affection and control, which they look to obtain when getting together to form a group. A *need for inclusion* relates to whether an individual tends to include others in interaction or wants to be recognised and included by others. A *need for affection* refers to how an individual relates to others and how much someone invests in a relationship or desires another to demonstrate affection towards him/her. Lastly, a *need for control* differentiates between people who desire to be dominated and guided, and those who prefer to lead and take responsibility.

RIASEC

The history of the RIASEC Occupational Interests framework dates back to a series of influential papers about vocational preferences by John Holland in the 1950s. His 1959 article in the Journal of Counselling Psychology titled "A Theory of Vocational Choice" is particularly striking because of the way he describes the choice of a vocation as an expression of one's personality. Holland's model is now often referred to as the RIASEC, as this is the acronym produced by his revised names for the six categories of occupations and orientations - *Realistic, Investigative, Artistic, Social, Enterprising, and Conventional*. In personality terms, the RIASEC model is generally construed as a dimensional, theory-based model. While the six categories are suggestive of a typological approach, Holland was clear in his views that the types are dimensional. Several researchers have previously reported on the associations between the RIASEC and other personality frameworks. DeFruyt and Mervielde (1999) reported significant (and substantial) correlations between each of the Big Five with at least one of the RIASEC scales. Openness, for example, was strongly positively correlated with the Artistic scale and negatively correlated with the Conventional scale. The Realistic and Investigative scales, according to DeFruyt and Mervielde, are not strongly correlated with any of the Big Five scales. For **WORK PROFILE**, the RIASEC model and research findings associated with it were considered as a framework as they help understand how an individual makes his/her choices, work satisfaction and performance, as well as career development.

Motivations, Aspirations and Values

The relationship between an individual's values and motivations, and work success has been long documented. Several theories point to the role of motivation and values to determine an individual's life goals as well as the decisions he/she makes, and to explain why individuals behave the way they do on a day-to-day basis.

Values are principles or standards of behaviour that govern one's judgement of what is important in life. Sociologist Morris Massey (1979) argued that values are formed during three periods: a) *The Imprint Period* – until the age of seven, our values are blindly absorbed from our parents; b) *The Modelling Period* – lasts between seven and approximately thirteen years of age, other sources such as teachers and role models influence our values and we try to imitate others; c) *The Socialization Period* – during adolescence and early adulthood (13 to 21 years), we are mainly influenced by our peers and actively seek others who are like us and share our values. Developed personal values therefore influence our attitudes and behaviours as they determine what we consider to be appropriate, right or wrong.

Values are also considered to be long-term objectives that can develop and adapt as a result of life experiences. Finegan (2000) proposed through his multidimensional approach that there are two sets of factors: *Humanity Factors* (such as courtesy, consideration and forgiveness) and *Vision Factors* (such as development, creativity and openness) which influence our decision-making and the way that we experience our surroundings. He suggested that an organisation that values both Vision and Humanity factors is likely to be perceived positively and therefore acquire a higher level of commitment from its employees that place importance on these values, as job applicants tend to look for companies that share the same values (Anderson, Fornell, & Mazvancheryl, 2004; Barnard, 1968; Chatman, 1989; Schneider, 1987).

Shalom Schwartz (2012) also conducted extensive research on this topic, and identified 10 cross cultural types of values: *Power, Achievement, Stimulation, Hedonism, Universalism, Benevolence, Tradition, Conformity, Security and Self-direction*.

Motivation on the other hand, is the desire or willingness to do something; the underlying reasons for an individual behaving in a particular way. Motivation influences emotional, social and cognitive forces that lead to behaviour. There are two types of motivation: extrinsic and intrinsic. *Extrinsic* motivation refers to factors external to the individual that influence their beliefs, attitudes, decisions and behaviours. In the context of work, extrinsic motivators may include salary, job perks/benefits, status and work conditions. *Intrinsic* motivation, on the other hand, refers to behaviour that is driven by internal rewards or feelings of personal satisfaction unique to an individual. In the context of work, intrinsic motivators include recognition, challenging work, purposeful work, achievement and opportunities for personal growth.

The relationship between employee values/motivations and job performance has been long established. Chung et al. (2008) suggested that work values are a set of durable beliefs in work. They reflect personal needs, job types or the environmental preferences of individuals. They also guide the work behaviour of individuals and can dictate their goals and job choices. Work values have an influence on hiring and retention rates, and can affect the willingness of individuals to work diligently (Boyatzis and Skelly, 1991). Chiu (2002) suggested that most individuals consider whether the image and values of an organization match their own in their job selections. If they match, it is an enhancer to the development of the individual and the organization; if not, it becomes an obstacle. It is imperative for companies to gain a deep understanding of the work values of their employees and ensure that the ownership of desirable work values is factored into the recruitment process (Chung et al., 2008).

Research on the relationship between employee motivation and job and organisational success is extensive. There is clear evidence establishing a link between motivation and employee job performance (Alonso & Lewis, 2001; Bright, 2007; Locke & Latham, 2002; Locke, Shaw, Saari, & Latham, 1981); retention (Rainall, 2004) and work attitudes (Kovach, 1995; Porter, Bigley, & Steers, 2003). With the recruitment and loss of talented employees associated with high financial and productivity costs, identifying motivated employees remain a primary objective of Human Resource practitioners.

Abraham Maslow (1943) proposes in his Theory of Human Motivation that needs can be classified in a pyramid to visualise the order in which they appear in an individual. He states that satisfaction of needs at lower levels elicits needs at the need level, and concludes that humans can only reach the top of the pyramid if lower level needs are satisfied. Figure 2 depicts Maslow's theory of motivation and hierarchy of needs.



Figure 1. Maslow's Theory of Motivation and Hierarchy of Needs.

Later, David McClelland proposed the "Needs Theory" of motivation. This is closely associated with learning concepts (Pardee, 1990) and is built on Abraham Maslow's work in his 1961 book; "The Achieving Society." The theory proposes that when a need is strong in a person, its effect is to motivate the person to behave in a way that will satisfy the need (Pardee, 1990). He identified three motivators that he believed we all have: *a need for achievement*, *a need for affiliation*, and *a need for power*. According to McClelland, these motivators are learned (which is why this theory is sometimes called the Learned Needs Theory). McClelland says that, regardless of our gender, culture, or age, we all have three motivating drivers, and one of these will be our dominant motivating driver and that this dominant motivator is largely dependent on our culture and life experiences.

WORK PROFILE considers motivations and values to be an important aspect in determining which tasks will keep an individual driven and satisfied, as well as the environment that he/she is likely to flourish in. Table 1 details the psychological theories and frameworks to which each factor of **WORK PROFILE** corresponds to.

Table 1. Work Profile Factors Mapped onto Other Theoretical Frameworks.

WORK PROFILE Factors	Corresponding Factor	Reference Model/Framework
Organisation	Conscientiousness	FFM, HEXACO
	Conventional	RIASEC
Compliance	Affiliation Need	Needs Theory
	Agreeableness	FFM, HEXACO
	Conventional	RIASEC

	Conformity, Tradition	Schwartz Theory of Basic Values
Dynamism	Achievement Need	Needs Theory
	Extraversion	FFM, HEXACO
	Enterprising	RIASEC
Sociability	Affiliation Need	Needs Theory
	Extraversion	FFM, HEXACO
	Need for Affection, Need for Inclusion	FIRO
	Social	RIASEC
	Benevolence, Conformity	Schwartz Theory of Basic Values
Sense of Responsibility	Social	RIASEC
	Conscientiousness	FFM, HEXACO
	Honesty-Humility	HEXACO
	Conformity, Tradition	Schwartz Theory of Basic Values
	Need for Control	FIRO
Persuasiveness	Power Need	Needs Theory
	Enterprising	RIASEC
	Extraversion, Agreeableness	FFM, HEXACO
	Need for Control	FIRO
Self-Control	Neuroticism	FFM
	Realistic	RIASEC
Tenacity	Achievement Need	Needs Theory
	Conscientiousness	FFM, HEXACO
	Conventional	RIASEC
Assertiveness	Enterprising	RIASEC
	Agreeableness, Conscientiousness, Neuroticism	FFM, HEXACO
Adaptability	Agreeableness, Openness	FFM, HEXACO
	Enterprising	RIASEC
Sense of Service	Agreeableness	FFM, HEXACO

	Social	RIASEC
	Openness	FFM, HEXACO
Ambition	Achievement Need	Needs Theory
	Conventional, Enterprising, Realistic	RIASEC
	Achievement	Schwartz Theory of Basic Values
	Achievement Need	Needs Theory
Autonomy	Social, Artistic, Realistic, Investigative	RIASEC
	Extraversion, Openness	FFM, HEXACO
	Self-Direction	Schwartz Theory of Basic Values
	Need for Control	FIRO
	Investigative, Realistic	RIASEC
Intellectual Curiosity	Openness	FFM, HEXACO
	Stimulation	Schwartz Theory of Basic Values
	Power Need	Needs Theory
	Agreeableness	FFM, HEXACO
Leadership	Power	Schwartz Theory of Basic Values
	Need for Control	FIRO
	Enterprising	RIASEC

2.2. Development of WORK PROFILE

WORK PROFILE was developed from a thorough, multistage process in a comprehensive and scientific manner. The development of the test started in 2015, as a result of accrued demand from companies using psychometric assessments in order to provide a “supplementary aid” in the recruitment process. Extensive empirical and theoretical research was conducted to create the dimensions, factors and items of **WORK PROFILE**. The questionnaire was also pretested several times to ensure that it met the criteria for reliability and validity established by internationally recognised psychological institutions. Special attention was also given to guard against social desirability as well as to simplify the ergonomics of the assessment.

2.2.1. Dimensions, Factors and Item Construction

WORK PROFILE was developed from a bank of items from the original questionnaire (existing only in French), the *Profil Compétences*. The original questionnaire then underwent review that resulted in **WORK PROFILE** which assesses three broad dimensions (*Workplace Traits, Aspirations and Values*, and *Social and Emotional Traits*), covering a total of 15 factors. The 15 factors are grouped into the 3

dimensions in order to provide a contemporary and accurate representation of the differences between the factors, and how they are assessed. Therefore, measuring these factors are the 90 questions, presented as paired statements (the equivalent of 180 items). Multiple reviews were conducted to ensure that the factors are clearly defined, and that items targeting each factor accurately reflect modern and business-relevant behaviours in lower level trades, thus resulting in the following:

- I. Greater differentiation between trait-based and motivation-based dimensions
- II. Inclusion of factors that target occupational and organisational fit
- III. Clearer item structure and wording
- IV. Updated use of terms and relevance of items to match the current working world
- V. Simple and easily comprehensible use of language

Thus, items for **WORK PROFILE** were written to promote ease of response as well as facilitate candidates' mental connection and understanding of it. As the questionnaire aimed to capture work behaviours in multiple contexts, items were checked for cultural bias, clarity, offensiveness and language accuracy.

The three dimensions and their relevant factors are measured as follows:

- I. **Workplace Traits** which include *Organisation, Compliance, Dynamism, Sociability* and *Sense of Responsibility*. This dimension and its factors are assessed as behavioural tendencies, and were selected based on research, as discussed in Chapter 2.1.1., that suggests that the traits measured are indicative predictors of job performance.
- II. **Social and Emotional Traits** which include *Persuasiveness, Self-Control, Tenacity, Assertiveness* and *Adaptability*. This dimension and its factors are assessed as behavioural orientations, and were chosen based on research, discussed in Chapter 2.1.1., that suggests that the traits measured are essential in order to manage social and emotional behaviours in the workplace.
- III. **Aspirations and Values** which include *Sense of Service, Ambition, Autonomy, Intellectual Curiosity* and *Leadership*. This dimension and its factors are assessed as preferences, and were selected based on research discussed in Chapter 2.1.1., relating to motivational predictors of employee success.

2.2.2. Ipsative Response Format

Response options for the 60 core questions of **WORK PROFILE** are written in binary, ipsative (forced-choice) format. In contrast to Likert scales, respondents choose between descriptors according to the extent of which the descriptor captures their preference. Ipsative formats have been found to reduce response bias (Chan, 2003; Saville & Willson, 1991; Ten Holt, van Duijn, & Boomsma, 2010). The fact that respondents cannot endorse all items eliminates acquiescence (the tendency to agree with all items on a questionnaire), reduces susceptibility to impression management and means that it is generally more difficult to fake (Brown & Maydeu-Olivares, 2011a). Ipsative formats make it impossible to provide the same response on all items, thereby resulting in greater differentiation of scores within a profile. Ipsative scales are also beneficial because there is no need to include negatively worded items in order to balance dimensions. Rather, mirrored items facilitate measurement of behaviours for different traits along the same domains. In Likert scales, it is necessary to include negatively worded items in order to balance responses and monitor acquiescence (Brown & Maydeu-Olivares, 2013; Edwards & Kenney, 1946). With ipsative formats, all items are worded positively yet both positive and negative behavioural indicators are captured.

Despite such advantages, ipsative questionnaires are underutilised in work-based assessments due to challenges for traditional normative scoring methodology, construct validation and interpretation (Cornwell & Dunlap, 1994; Loo, 1999; Meade, 2004).

In an ipsative test, the respondent must make a comparative judgement between two or more items to give an answer. Conversely, in a normative test, judgement is viewed in absolute terms for each evaluated personality trait and preferences are not compared between different traits. Classical Test Theory (CTT) scoring does not reflect the psychological process of comparison, as it ignores the fact that respondents indicate a preference, not an absolute judgement. As a consequence, ipsative test scores that are generated from CTT scoring approaches might be distorted. The potential solution therefore is to keep the ipsative format but change the way the scores are calculated, in order to reflect realistic comparative judgements (see Chapter 2.2.4.).

2.2.3. Normative Response Format

On the other hand, response options for the 30 normative items follow a format differing from the ipsative response format, meaning that rather than choosing between two items from differing factors, the respondent must choose between one item representing a particular factor, and another representing the behavioural opposite of that factor (scored as a lack of the factor). The combination of normative scoring with an ipsative means that the measurement of factors is fairer; aside from choosing preferred factors over others, the candidate is also scored directly on each factor.

2.2.4. Thurstonian Item Response Theory (IRT) and Ipsative Responses

To solve the problem of ipsative scoring, researchers relied on the work of Thurstone (1927, 1931) who proposed the law of comparative judgements, which assigns the result of a comparative judgment to the "relative utility" of objects compared for each individual. Applied to personality assessment, the notion of utility refers to the degree to which there is a match between the item and the personality trait of the respondent. Therefore, an individual will choose the item that corresponds to them the most.

Despite its potential benefits, Thurstonian models were previously not utilised in ipsative test as (Brown, 2010):

- I. There was no efficient estimation method available.
- II. Computer technology was not powerful enough to facilitate analysis.
- III. Previous statistical programmes did not facilitate the calculation of scores because the variance of error for each question is 0.

Recent changes were made to the Thurstonian second order model to produce the Thurstonian IRT model (Brown & Maydeu-Olivares, 2011, 2013). This model bypasses the notion of latent utility, and directly connects the choice made by the respondents to the latent personality traits measured by the test. This change simplifies the model and allows for it to be applied to much longer questionnaires.

Brown & Maydeu-Olivares (2011, 2013) showed that when applied to an ipsative questionnaire with blocks of four items, this method allows not only to correct the distortion of the scores, but also to recover the levels of "absolute" personality traits from responses based on comparative judgements. The scores provided were similar to those of a normative test.

The nature of this scoring procedure, based on the Maximum a posteriori estimator (MAP), was completely different from that of the CTT, which "adds" the dimension points for each selected item.

MAP is an estimation procedure which has the goal to find a parameter value θ (the value on the latent trait continuum) that best explains the data (the pattern of binary outcome responses) by assuming a joint distribution on both, the data and the θ , with θ governing the distribution on the data (i.e. the posterior distribution). More simply, with the Maximum a posteriori estimator, the overall response pattern of the test is taken into account for the calculation of the score. By comparing the choice and preference indicated by the respondent while taking into account the specific characteristics of the items, the method is able to deduce (or recover) their absolute level of personality trait for each dimension.

For **WORK PROFILE**, Thurstonian IRT was used to revise, rewrite and validate the items and factors. Figure 3 displays the model of Thurstonian IRT used to analyse items and factors in the questionnaire.

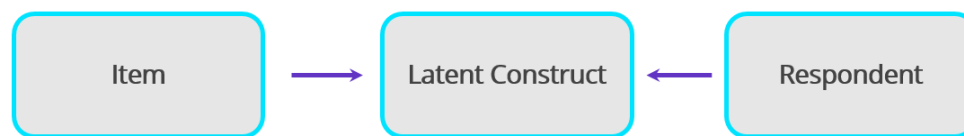


Figure 3. Model of Thurstonian IRT analyses utilised in the development of Work Profile.

2.2.4.a. Item Facility

Item facility refers to an item's ability to differentiate between respondents and identify floor or ceiling effects. An item has floor or ceiling effects when the probability of endorsing a response option is much lower or greater than the probability of endorsing the alternate option. As items for **WORK PROFILE** were binary, item facility was examined from the proportion of responses for each option. The more equivalent the proportion, the more balanced the item. For **WORK PROFILE**, items with floor or ceiling effects were discarded or modified. The facility of changed or new items was then examined with subsequent pretests.

2.2.4.b. Item Difficulty and Discrimination

Item difficulty is the first parameter of an IRT model. It is defined as the location of the item on the ability continuum, where the probability of endorsing the item is equal to 50% of the participants. In a personality assessment, the difficulty parameter reflects the level of a trait needed for participants to have a 50% chance of choosing the item. An item is said to be "easy" when the probability that 50% of individuals who endorsed it only have a low level on the trait measured by the item (e.g. an "easy" item measuring Dynamism might be chosen by 50% of participants having only a low level of trait for Dynamism). On the contrary, an item is said to be "difficult", when the probability that 50% of individuals who endorsed it have a high level of trait (e.g. a "difficult" item measuring Dynamism might reach the probability of being chosen by 50% of the participants only if they have a high level of trait for Dynamism).

Item difficulty can be visualised on the Item Characteristic Curve (ICC). The ICC is an S-shaped curve that represents the relationship between ability levels of the latent trait and the probability of endorsing the item measuring it. It is obtained by modelling the probability of endorsing the items using a logistic model.

Item discrimination is the second parameter of the IRT model and refers to the shape of the curve. It is a measure of an item's ability to distinguish between respondents on the target construct. A high

discrimination parameter value reflects a high ability to differentiate between participants; the probability of choosing the correct answer (or endorsing the item for a personality assessment) increases more rapidly as the ability (or level of trait) increases.

As the Thurstonian IRT model is conceptualised from a factor analysis perspective, item difficulty is referred to as "threshold" instead of "difficulty" and item discrimination parameter as "factor loadings". Factor loadings and threshold were examined for all Items. Items that were too difficult, too easy or did not discriminate enough were modified or removed.

2.2.4.c. Item and Factor Reliability

Item reliability is the degree to which items on a questionnaire have no error. The reliability of an item is determined by the internal consistency of the item and its relationship with other items that measure the same factor. Internal reliability was measured using composite reliability generated from item factor loadings. Composite reliability ranges from 0 to 1 with a recommended coefficient of 0.7. For **WORK PROFILE**, items with the smallest factor loadings were changed or deleted. Modified questionnaires were then trialled with successive pretests.

2.2.5. Pretests and Pilot Study

Pretests were conducted with a range of male and female professionals across the globe to ensure items were accurate, relevant, clear, concise, culture-free, and bias-free and that they reflected targeted work behaviours. After five pretests, the final questionnaire was piloted with 373 English-speaking professionals from 28 countries. Pretest samples comprised of English speakers from different continents in order to establish the questionnaire as an international measure of work personality and motivation.

WORK PROFILE was developed to ensure items were culturally equivalent from the early stages of development, rather than focusing on the cultural adaption after the questionnaire validity had been established within a single context – the typical practice of test developers.

Composite reliabilities for each factor were calculated from generated item factor loadings. Factors were flagged if reliabilities were less than 0.55. Also, factors that did not converge were inspected for poorly performing items and were then reviewed, modified or changed and then subjected to another round of pretest. The overall scale composite reliability for the Pilot study was 0.78 (see Chapter 5).

3. Scoring

A raw score is the total number of responses endorsed by a candidate for each factor measured. In the context of a recruitment personality assessment, raw scores mean little as they do not provide effective or appropriate comparisons between candidates across a specific interest. Norms enable Human Resource practitioners, managers and recruiters to evaluate an individual's score relative to the score of other delegates who have completed the same assessment.

3.1. The Relationship between Raw Scores and Norms

The best norm group is one that is most representative of the characteristics of the individuals taking a test. Applied assessment assumes personality traits, interests, motivations, values, attitudes and perceptions are distributed along a normal probability or bell-shaped curve. Therefore, norm groups effectively identify individual scores across this distribution. In any given population, levels of a target attribute may be found. Therefore, for a given attribute or trait, the normal curve represents the norm or average performance of a population and the scores that are above and below the average for that population (Kaplan & Saccuzzo, 2008). For norms to be meaningful, the sample used must reflect the intended population of test takers. If a test is developed for engineers, the norm sample must consist of engineers in the population. Similarly, if an assessment was developed for nurses, it is not appropriate for use in the selection of chemists. Norm groups can range in size from a few hundred to a hundred thousand. However, the larger the norm group sample size, the closer the approximation to a normal population distribution. Figure 4 shows the normal bell curve.

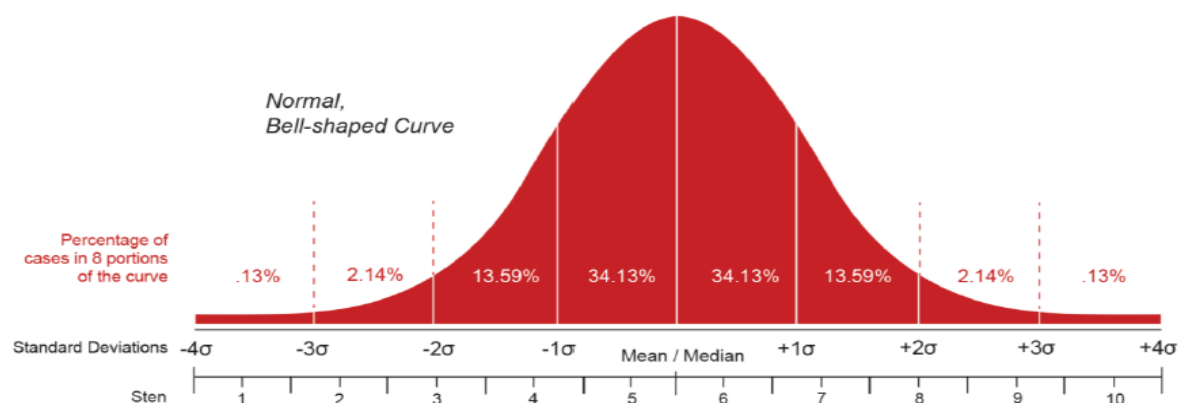


Figure 3. Normal Bell-shaped curve showing the distribution of percentages, standard deviations and standard ten (Sten).

Key to the creation of norms is the conversion of raw scores into derived scores. This process provides a system of common metrics that readily facilitates the interpretation of test scores. For **WORK PROFILE**, the forms of score standardization used were: 1) Mean; 2) Median; 3) Standard deviation; 4) Standard score; 5) Standard ten (sten); and 6) Standard error of measurement. Table 2 describes the various types of derived scores used in the validation of **WORK PROFILE**.

Table 2. Types of Derived Scores used to Standardise Work Profile.

Sr. No.	Derived Score	Symbol	Definition and Characteristics
1	Mean	μ	Mean is the average score of all participants who complete an assessment. The mean is calculated by dividing the sum of all scores by the number of participants. The formula for calculating the Mean is: $\mu = (\Sigma X) / N,$ where ΣX is the sum of all the population observations, and N is the number of population observations.
2	Median	$\sim x$	Median is the score that separates the highest scores from the lowest. The median can be found by arranging all scores in order from lowest to highest and selecting the score in the middle. Like Mean, Median is a measure of the general tendency of a group of scores.
3	Standard Deviation	Σ	Standard Deviation is a measure of the amount of variation across scores for candidates who take an assessment. A small standard deviation indicates scores are close to the mean, while a large value tells us there is much variability across individuals. The formula for calculating Standard Deviation is: $\sigma = \text{sqrt} [\Sigma (X_i - \mu)^2 / N],$ where μ is the population mean, X_i is the ith element from the population, and N is the number of elements in the population.
4	Standard Score	Z	Standard Score is an estimate of whether an individual's score is above average, average or below average compared to peers. Scored on a scale from -4 to 4 with zero in the middle representing average. Positive scores are above average while negative scores are below average. Z scores are calculated using the formula: $z = (x - \mu_x) / \sigma_x,$ where X is a normal random variable, μ is the mean of X, and σ is the standard deviation of X.
5	Standard Ten	Sten	For a Standard Score with scores ranging from 1 – 10, here Mean is 5.5 and Standard Deviation equals 2. The formula for calculating Sten Score is: $\text{Sten} = (X - \mu / \sigma)^2 + 5.5,$ where X is the candidate's raw score, μ is the mean of raw scores for all candidates, σ is the standard deviation of raw scores on the standardisation sample, multiplied by 2 + 5.5.

Sten Scores are rounded to the nearest whole number with values less than 1 or greater than 10 rounded up or down respectively.		
6	Standard Error of Measurement	SEM SEM is related to test reliability and provides an indication of the dispersion of measurement errors associated with test scores from a group of candidates. It helps us to identify level of errors when estimating candidate's true scores from their observed scores. The smaller the SEM, the more accurate are the assessments that are being made. The formula for calculating SEM is: $SEM = \sigma \times \sqrt{1 - \text{reliability}}.$

3.2. Transforming Raw Scores to Standardised Scores

For the 90 items in **WORK PROFILE**, the minimum and maximum overall score for each attribute is 0 and 10. The conversion of raw scores to standardised scores is achieved using a four step process, as below.

Step 1 Ipsative scores are calculated for each scale by summing all forced-choice responses for each attribute. For each response, 1 is assigned to the response that is endorsed and 0 for the response that is not endorsed.

Step 2 Scores are repositioned on a scale from 0 to 10 by using the following formula:

$$\text{score} = \left(\frac{\text{raw score} - (\text{min})}{(\text{max}) - (\text{min})} \right) * 10$$

Whereby max refers to the maximum possible score of each scale and min refers to the minimum possible score of each scale.

Step 3 To enable the comparisons between individuals, scores for each scale are standardised using the mean and standard deviation across respondents. This process produces a z score using the formula:

$$z = (\text{score} - (\mu/\sqrt{n}))$$

Step 4 Standardised scores are then transformed into sten scores from 1 to 10 using the formula:

$$\text{sten} = z * 2 + 5.5$$

3.3. Standard Error of Measurement

No test score is perfect. Therefore, we need an estimate of the possible amount of error present within a score. The amount of error associated with a score is known as the standard error of measurement (SEM). SEM helps us to understand what is a candidate's true score and the extent to which it differs from the obtained score (Harvill, 1991). In an ideal world, the SEM would be calculated by having an individual take an assessment on many occasions over a long period of time. The extent to which there is variation across each score obtained by the candidate would be classified as the standard error of

measurement. However, as it is not feasible to assess candidates multiple times, the SEM must be calculated using a different approach.

The SEM is an estimate that is added to or subtracted from a candidate's score in order to establish a level of confidence about a score. From this confidence band, we are able to determine the extent to which a score may vary. Standard error of measurement is related to reliability. Therefore, as reliability increases, SEM decreases. For a series of scores, approximately 68% of the time, an observed score will lie within +1.0 and -1.0 SEM of the true score. On 95% of occasions, the observed score will be between +1.96 and -1.96 SEM of the true score. Whilst 99% of the time, the observed score will lie within +2.58 and -2.58 SEM of the true score. The 95% confidence band of a score is calculated using the formula:

$$\text{Confidence band} = \text{score} \pm 1.96 \times \text{SEM}$$

Therefore, for **WORK PROFILE**, a factor with a score of 5 and SEM of 0.2, has a confidence band of 4.608 to 5.392. This means we can be 95% confident that an individual's true score lies in this range.

3.4. Choosing an Appropriate Norm Group

Norms make it possible to give meaning to an individual's score. The interpretation of the same score may therefore differ depending on individual differences, such as study level. For accurate interpretation, it is necessary to compare an individual's score with a group that is most similar to him/her. To choose the appropriate norm, first identify the country of the candidate. Then from that country, identify the age, gender, level of education, number of years' work experience, socio-professional status, industry or occupational group that is most appropriate for the candidate. Alternatively, it is possible to create your own internal norms, based on the candidates that have already been assessed.

3.5. Creating In-house Norms

It is possible to generate internal organisational norms if there is a relevant sample of 50 or more individuals. In-house norms ensure that comparisons between individuals and the reference group reflect the target role, tasks and work environment as closely as possible. In-house norms are particularly valuable in situations where standard norms are not available, roles are new, the organisation has a unique client base or business function, and in cases where the applicant pool for a role is significantly higher or lower than expected.

To create in-house norms, first calculate the mean and standard deviation for each factor on **WORK PROFILE** and score the results based on the step-by-step guide outlined in Chapter 3.2. Organisations will then be able to compare individual applicants or employees to a local normal group.

3.6. Link between Scores and Report Comments

Comments on **WORK PROFILE** report reflect the scoring process. An individual receives a score of 0 to 10 (as demonstrated in Chapter 3.2.). The comments provide an extensive interpretation of all the factors that are within the three dimensions of the test. They are chosen from six levels according to the score of the candidate, and detail the individual's attributes according to his/her score. Higher scores are accompanied by stronger descriptors, weaker scores are accompanied by weaker

descriptors, and so on. For example, a candidate with a low score on a factor may be described as 'not appearing to have a tendency/preference' for the relevant factor. A high score on the other hand may describe an individual as having 'a strong tendency/preference' for the relevant factor. Scores should be interpreted in relative terms as the order of the targeted profiles is more important than their absolute score.

4. Sample Descriptives

The pilot study sample comprised of 373 English speaking participants (196 females and 177 males), from a large range of occupational groups. Data was collected from participants using the professional research and sampling firm Amazon Mechanical Turk (303 participants), as well as the Research and Innovation Testing Centre of *Central Test* (70 participants). To ensure the quality of data obtained, participants whose answers were deemed unreliable or who completed the questionnaire in less than 6 minutes, were excluded from the analyses. Tables 3 to 5 provide an overview of the sample descriptives.

Table 3. Descriptives for Gender.

Gender	Percentage (%)	N
Female	53	196
Male	47	177

Table 4. Descriptives for Age Groups.

Age Group	Percentage (%)	N
Not defined	4%	15
18-25	14%	52
26-33	38%	141
34-41	20%	74
42-49	11%	42
50-56	8%	28
57+	6%	21

Table 5. Descriptives for Candidate Country of Origin.

Candidate Country	Percentage (%)	N
India	26	96
USA	57.10	213
Others	17.16	64

5. Norms

Norms help us to compare an individual's score on a test with the scores of others who completed the same assessment, usually under similar conditions. These individuals are representative of the population for which an assessment is intended, and are therefore referred to as the norm group. The closer the match between the characteristics of the sample and the target population, the more accurate the distribution will be as a ranking guide. With norms, test developers and psychologists are able to establish comparisons between individuals and make more objective decisions. Norm groups enable us to attribute meaning to test scores and quantify an individual's performance, level of an attribute or trait relative to his/her peers.

In the case of work-based assessments, norm groups help Human Resource Managers, recruiters and coaches to:

- I. Determine a candidate's suitability for a role relative to the job specification.
- II. Evaluate a candidate's abilities across a range of desired competencies.
- III. Identify an employee's skill deficits and tailor training for development.
- IV. Identify a candidate's potential team and organisational culture fit.
- V. Determine how a candidate is likely to respond to stress or changes in the work environment.
- VI. Establish objective metrics of performance.

Tables 6 to 9 detail the mean, standard deviation, median, minimum, maximum, range and standard estimate scores for the English-speaking sample on factors of **WORK PROFILE**. These tables are divided into categories based on age, gender and country.

Table 6. Factor Score Norms for the Whole Sample.

Factor	Mean	SD	Median	Min	Max	Range	SE
Adaptability	4.78	2.15	4.60	0.00	9.20	9.20	0.11
Ambition	5.50	2.32	5.80	0.80	10.00	9.20	0.12
Assertiveness	5.55	2.71	5.80	0.00	10.00	10.00	0.14
Autonomy	5.73	2.39	5.80	0.00	10.00	10.00	0.12
Compliance	5.45	2.48	5.80	0.00	10.00	10.00	0.13
Dynamism	5.11	2.39	5.00	0.00	10.00	10.00	0.12
Intellectual Curiosity	6.45	2.38	6.70	0.00	10.00	10.00	0.12
Leadership	3.86	2.83	4.20	0.00	10.00	10.00	0.15
Organisation	5.75	2.56	5.80	0.00	10.00	10.00	0.13
Persuasiveness	3.95	2.32	4.20	0.00	10.00	10.00	0.12

Self-Control	5.78	2.45	5.80	0.00	10.00	10.00	0.13
Sense of Responsibility	4.22	2.48	4.20	0.00	10.00	10.00	0.13
Sense of Service	4.47	2.78	4.20	0.00	10.00	10.00	0.14
Sociability	5.17	3.10	5.00	0.00	10.00	10.00	0.16
Tenacity	5.27	2.10	5.00	0.00	10.00	10.00	0.11

Table 7. Factor Score Norms for Age Groups.

Factor	Mean	SD	Median	Min	Max	Range	SE
18 – 25							
Adaptability	4.63	2.38	4.60	0.00	9.20	9.20	0.33
Ambition	5.76	2.20	5.80	1.70	9.20	7.50	0.31
Assertiveness	5.07	2.66	4.60	0.80	9.20	8.40	0.37
Autonomy	5.40	2.13	5.80	1.70	9.20	7.50	0.30
Compliance	5.20	2.19	5.00	1.70	10.00	8.30	0.30
Dynamism	5.25	2.30	5.00	0.80	9.20	8.40	0.32
Intellectual Curiosity	7.15	1.96	7.50	2.50	10.00	7.50	0.27
Leadership	3.40	2.80	3.30	0.00	10.00	10.00	0.39
Organisation	5.17	2.56	5.00	0.00	9.20	9.20	0.36
Persuasiveness	4.20	2.41	4.60	0.00	9.20	9.20	0.33
Self-Control	5.81	2.78	5.80	0.00	10.00	10.00	0.39
Sense of Responsibility	4.27	2.81	4.20	0.00	10.00	10.00	0.39
Sense of Service	4.42	2.95	4.20	0.00	10.00	10.00	0.41
Sociability	5.23	3.28	5.40	0.00	10.00	10.00	0.45
Tenacity	5.41	2.08	5.80	0.80	9.20	8.40	0.29
Factor	Mean	SD	Median	Min	Max	Range	SE
26 – 33							
Adaptability	4.74	2.03	4.60	0.00	9.20	9.20	0.17
Ambition	5.83	2.16	5.80	0.80	10.00	9.20	0.18

Assertiveness	5.46	2.52	5.80	0.00	10.00	10.00	0.21
Autonomy	5.42	2.35	5.80	0.00	10.00	10.00	0.20
Compliance	5.38	2.35	5.80	0.00	10.00	10.00	0.20
Dynamism	5.13	2.30	5.00	0.00	10.00	10.00	0.19
Intellectual Curiosity	6.24	2.46	6.70	0.00	10.00	10.00	0.21
Leadership	4.12	2.64	4.20	0.00	9.20	9.20	0.22
Organisation	5.85	2.21	5.80	0.00	10.00	10.00	0.19
Persuasiveness	4.06	2.29	4.20	0.00	10.00	10.00	0.19
Self-Control	5.90	2.29	5.80	0.00	10.00	10.00	0.19
Sense of Responsibility	4.48	2.27	4.20	0.00	10.00	10.00	0.19
Sense of Service	4.55	2.59	4.20	0.00	10.00	10.00	0.22
Sociability	5.12	2.86	5.00	0.00	10.00	10.00	0.24
Tenacity	5.05	1.94	5.00	0.00	10.00	10.00	0.16

Factor	Mean	SD	Median	Min	Max	Range	SE
34 – 41							
Adaptability	4.33	2.16	4.60	0.00	8.50	8.50	0.25
Ambition	5.46	2.59	5.40	0.80	10.00	9.20	0.30
Assertiveness	5.85	2.82	5.80	0.00	10.00	10.00	0.33
Autonomy	6.06	2.44	6.70	0.80	10.00	9.20	0.28
Compliance	5.52	2.53	5.80	0.00	10.00	10.00	0.29
Dynamism	4.96	2.55	5.00	0.00	10.00	10.00	0.30
Intellectual Curiosity	6.56	2.52	6.70	0.00	10.00	10.00	0.29
Leadership	3.75	2.89	4.20	0.00	9.20	9.20	0.34
Organisation	6.02	2.90	6.70	0.00	10.00	10.00	0.34
Persuasiveness	4.14	2.34	4.20	0.00	10.00	10.00	0.27
Self-Control	5.57	2.55	5.80	0.80	10.00	9.20	0.30
Sense of Responsibility	3.85	2.61	3.30	0.00	10.00	10.00	0.30
Sense of Service	4.14	2.76	4.20	0.00	10.00	10.00	0.32
Sociability	5.22	3.46	5.40	0.00	10.00	10.00	0.40

Tenacity	5.15	2.14	5.00	0.00	10.00	10.00	0.25
-----------------	------	------	------	------	-------	-------	------

Factor	Mean	SD	Median	Min	Max	Range	SE
---------------	-------------	-----------	---------------	------------	------------	--------------	-----------

42 – 49

Adaptability	4.90	2.48	4.60	0.80	8.50	7.70	0.38
Ambition	4.95	2.76	5.00	0.80	10.00	9.20	0.43
Assertiveness	5.72	2.90	6.70	0.00	10.00	10.00	0.45
Autonomy	6.47	2.39	6.70	1.70	10.00	8.30	0.37
Compliance	5.70	2.66	6.70	0.00	10.00	10.00	0.41
Dynamism	4.86	2.63	4.20	0.00	10.00	10.00	0.41
Intellectual Curiosity	6.23	2.23	6.70	0.80	10.00	9.20	0.34
Leadership	3.31	2.80	2.50	0.00	9.20	9.20	0.43
Organisation	6.22	2.64	6.70	0.00	10.00	10.00	0.41
Persuasiveness	4.05	2.26	3.75	0.00	9.20	9.20	0.35
Self-Control	6.28	2.50	6.70	0.80	10.00	9.20	0.39
Sense of Responsibility	4.17	2.23	4.20	0.00	9.20	9.20	0.34
Sense of Service	4.45	3.03	4.20	0.00	10.00	10.00	0.47
Sociability	4.68	3.07	4.20	0.00	10.00	10.00	0.47
Tenacity	5.00	2.05	5.00	0.80	9.20	8.40	0.32

Factor	Mean	SD	Median	Min	Max	Range	SE
---------------	-------------	-----------	---------------	------------	------------	--------------	-----------

50 – 56

Adaptability	5.11	1.72	4.60	2.30	9.20	6.90	0.32
Ambition	4.41	1.79	4.60	0.80	7.50	6.70	0.34
Assertiveness	6.31	2.67	7.50	0.80	10.00	9.20	0.50
Autonomy	6.14	2.28	6.70	0.00	9.20	9.20	0.43
Compliance	5.84	3.14	6.70	0.00	10.00	10.00	0.59
Dynamism	5.21	2.36	4.60	0.80	10.00	9.20	0.45
Intellectual Curiosity	6.58	2.39	6.70	1.70	10.00	8.30	0.45

Leadership	4.05	3.39	4.20	0.00	9.20	9.20	0.64
Organisation	5.65	2.98	5.80	0.00	10.00	10.00	0.56
Persuasiveness	2.71	2.30	2.50	0.00	9.20	9.20	0.43
Self-Control	5.15	2.49	5.80	0.00	9.20	9.20	0.47
Sense of Responsibility	3.99	2.96	3.30	0.00	10.00	10.00	0.56
Sense of Service	4.72	3.11	3.75	0.00	10.00	10.00	0.59
Sociability	5.09	3.29	5.80	0.00	10.00	10.00	0.62
Tenacity	6.37	2.37	6.70	0.00	10.00	10.00	0.45

Factor	Mean	SD	Median	Min	Max	Range	SE
57+							
Adaptability	5.50	2.10	5.40	0.80	8.50	7.70	0.46
Ambition	4.88	2.40	5.00	0.80	10.00	9.20	0.52
Assertiveness	6.14	3.11	7.50	0.80	10.00	9.20	0.68
Autonomy	6.15	2.41	5.80	1.70	10.00	8.30	0.53
Compliance	4.76	3.18	5.00	0.00	9.20	9.20	0.69
Dynamism	5.36	2.13	5.80	1.70	8.30	6.60	0.47
Intellectual Curiosity	5.87	2.62	6.70	0.00	10.00	10.00	0.57
Leadership	4.05	3.71	4.20	0.00	9.20	9.20	0.81
Organisation	5.20	3.09	5.00	0.80	10.00	9.20	0.67
Persuasiveness	3.65	2.38	4.20	0.00	7.50	7.50	0.52
Self-Control	5.39	2.57	5.80	0.00	9.20	9.20	0.56
Sense of Responsibility	4.01	2.83	4.20	0.00	10.00	10.00	0.62
Sense of Service	5.24	2.87	5.80	0.00	9.20	9.20	0.63
Sociability	6.30	2.98	7.50	0.00	10.00	10.00	0.65
Tenacity	5.52	2.28	6.70	0.80	8.30	7.50	0.50

Table 8. Factor Score Norms for Gender.

Factor	Mean	SD	Median	Min	Max	Range	SE
<i>Females</i>							
Adaptability	4.78	2.36	4.60	0.00	9.20	9.20	0.17
Ambition	5.28	2.33	5.00	0.80	10.00	9.20	0.17
Assertiveness	5.67	2.84	5.80	0.00	10.00	10.00	0.20
Autonomy	5.83	2.42	5.80	0.00	10.00	10.00	0.17
Compliance	5.51	2.54	5.80	0.00	10.00	10.00	0.18
Dynamism	5.21	2.51	5.00	0.00	10.00	10.00	0.18
Intellectual Curiosity	6.20	2.39	6.70	0.00	10.00	10.00	0.17
Leadership	3.43	2.83	3.30	0.00	9.20	9.20	0.20
Organisation	5.70	2.68	5.80	0.00	10.00	10.00	0.19
Persuasiveness	3.78	2.28	3.30	0.00	10.00	10.00	0.16
Self-Control	5.63	2.52	5.80	0.00	10.00	10.00	0.18
Sense of Responsibility	4.39	2.53	4.20	0.00	10.00	10.00	0.18
Sense of Service	4.98	2.92	4.20	0.00	10.00	10.00	0.21
Sociability	5.30	3.17	5.80	0.00	10.00	10.00	0.23
Tenacity	5.23	2.18	5.80	0.00	10.00	10.00	0.16
<i>Males</i>							
Adaptability	4.79	1.91	4.60	0.00	9.20	9.20	0.14
Ambition	5.74	2.29	5.80	0.80	10.00	9.20	0.17
Assertiveness	5.41	2.56	5.00	0.00	10.00	10.00	0.19
Autonomy	5.62	2.36	5.80	0.00	10.00	10.00	0.18
Compliance	5.39	2.41	5.80	0.00	10.00	10.00	0.18
Dynamism	5.00	2.24	5.00	0.00	10.00	10.00	0.17
Intellectual Curiosity	6.72	2.34	6.70	0.00	10.00	10.00	0.18
Leadership	4.33	2.76	4.20	0.00	10.00	10.00	0.21

Organisation	5.81	2.44	5.80	0.00	10.00	10.00	0.18
Persuasiveness	4.14	2.36	4.20	0.00	9.20	9.20	0.18
Self-Control	5.94	2.38	5.80	0.00	10.00	10.00	0.18
Sense of Responsibility	4.04	2.42	4.20	0.00	10.00	10.00	0.18
Sense of Service	3.90	2.51	4.20	0.00	10.00	10.00	0.19
Sociability	5.03	3.03	5.00	0.00	10.00	10.00	0.23
Tenacity	5.32	2.00	5.00	0.00	10.00	10.00	0.15

Table 9. Factor Score Norms by Country.

Factor	Mean	SD	Median	Min	Max	Range	SE
India							
Adaptability	4.81	1.94	5.40	0.00	9.20	9.20	0.20
Ambition	5.59	1.80	5.80	0.80	10.00	9.20	0.18
Assertiveness	5.37	2.17	5.00	0.00	10.00	10.00	0.22
Autonomy	4.92	2.18	5.00	0.00	10.00	10.00	0.22
Compliance	5.68	1.87	5.80	0.00	10.00	10.00	0.19
Dynamism	5.04	2.10	5.00	0.00	10.00	10.00	0.21
Intellectual Curiosity	7.05	1.97	7.50	0.00	10.00	10.00	0.20
Leadership	4.43	2.19	4.20	0.00	9.20	9.20	0.22
Organisation	5.85	2.18	5.80	0.00	10.00	10.00	0.22
Persuasiveness	4.48	2.29	4.20	0.00	9.20	9.20	0.23
Self-Control	5.86	2.04	5.80	0.00	10.00	10.00	0.21
Sense of Responsibility	4.18	1.86	4.20	0.00	10.00	10.00	0.19
Sense of Service	4.41	2.16	4.20	0.00	10.00	10.00	0.22
Sociability	5.95	2.45	6.70	0.00	10.00	10.00	0.25
Tenacity	4.74	1.92	4.20	0.00	9.20	9.20	0.20

Factor	Mean	SD	Median	Min	Max	Range	SE
USA							

Adaptability	4.59	2.22	4.60	0.00	9.20	9.20	0.15
Ambition	5.37	2.48	5.00	0.80	10.00	9.20	0.17
Assertiveness	5.64	2.92	5.80	0.00	10.00	10.00	0.20
Autonomy	6.21	2.40	6.70	0.00	10.00	10.00	0.16
Compliance	5.76	2.44	5.80	0.00	10.00	10.00	0.17
Dynamism	5.26	2.46	5.00	0.00	10.00	10.00	0.17
Intellectual Curiosity	6.17	2.59	6.70	0.00	10.00	10.00	0.18
Leadership	3.41	3.01	3.30	0.00	9.20	9.20	0.21
Organisation	5.90	2.68	5.80	0.00	10.00	10.00	0.18
Persuasiveness	3.64	2.29	3.30	0.00	10.00	10.00	0.16
Self-Control	5.82	2.59	5.80	0.00	10.00	10.00	0.18
Sense of Responsibility	3.99	2.60	4.20	0.00	10.00	10.00	0.18
Sense of Service	4.40	2.96	4.20	0.00	10.00	10.00	0.20
Sociability	4.68	3.21	4.20	0.00	10.00	10.00	0.22
Tenacity	5.52	2.03	5.80	0.00	10.00	10.00	0.14

Factor	Mean	SD	Median	Min	Max	Range	SE
<i>Other</i>							
Adaptability	5.39	2.18	5.40	0.80	9.20	8.40	0.27
Ambition	5.80	2.49	5.80	0.80	10.00	9.20	0.31
Assertiveness	5.49	2.74	5.80	0.00	10.00	10.00	0.34
Autonomy	5.37	2.31	5.80	1.70	10.00	8.30	0.29
Compliance	4.06	2.93	3.30	0.00	10.00	10.00	0.37
Dynamism	4.70	2.51	4.20	0.00	10.00	10.00	0.31
Intellectual Curiosity	6.46	2.07	6.70	0.80	10.00	9.20	0.26
Leadership	4.48	2.83	4.20	0.00	10.00	10.00	0.35
Organisation	5.09	2.63	5.00	0.00	9.20	9.20	0.33
Persuasiveness	4.19	2.35	3.75	0.00	10.00	10.00	0.29
Self-Control	5.52	2.59	5.80	0.00	10.00	10.00	0.32
Sense of Responsibility	5.05	2.73	5.00	0.00	10.00	10.00	0.34

Sense of Service	4.78	3.02	4.60	0.00	10.00	10.00	0.38
Sociability	5.65	3.35	6.25	0.00	10.00	10.00	0.42
Tenacity	5.27	2.43	5.00	0.80	9.20	8.40	0.30

6. Reliability

Reliability refers to the extent to which an assessment is free from error and the degree to which items on a scale are measuring the same construct. For a measure to be reliable, the same results must be consistently generated over time. Reliability is important for an assessment as it enables us to have confidence that scores on a test are really typical of an individual's behaviour or attitude. A primary concern, when developing and using tests, is the extent to which random measurement errors influence test performance. If a test is unreliable, it is impossible for it to be valid as accuracy of results (validity) depends on the consistency with which the same results are produced. High reliability of an assessment helps to guard against biases that may be caused from errors in human judgment during the recruitment and selection process.

6.1. Internal Reliability

Internal reliability is a measure of the degree to which items on a scale are internally consistent and measure the same global construct or dimension of a construct. Composite reliability is a robust technique used to assess the internal consistency of items on a scale. It is mathematically defined as the amount of scale score variance accounted for by all underlying factors. Cronbach alpha, while a popular way of determining internal reliability, has a major weakness in that it assumes all items on a scale are equally important and therefore weighs all items similarly. As a consequence, if there is an increase in the number of items, reliability would artificially increase even though loadings may be low and not significant.

Therefore, composite reliability is a recommended alternative as it makes use of both item loadings and measurement error (Dunn, Baguley, & Brunsden, 2014; Raykov, 1997; Schmitt, 1996). Internal reliability values vary from 0 to 1, with higher values being more desirable. As a rule of thumb, a reliability of 0.70 or higher is recommended (Nunnally, 1978), with values greater than 0.9 (excellent); 0.7 to 0.9 (good); 0.6 to 0.7 (acceptable); 0.5 to 0.6 (poor) and values less than 0.5 (unacceptable). Composite reliability should be ≥ 0.70 .

The formula for Composite Reliability (CR) is:

$$CR = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + (\sum \epsilon_i)}$$

The composite reliability for **WORK PROFILE** was computed on the English-speaking sample ($n = 373$), and can be seen in Table 10. All the factors showed a composite reliability of above .6, with *Leadership* and *Sociability* at above .9. **WORK PROFILE** also has an overall composite reliability of .78, indicating that the scale has a very good internal consistency and is highly reliable.

Table 10. Composite Reliability for All Factors.

Factor	Composite Reliability
Adaptability	.83
Ambition	.79

Assertiveness	.84
Autonomy	.71
Compliance	.78
Dynamism	.71
Intellectual Curiosity	.80
Leadership	.91
Organisation	.79
Persuasiveness	.65
Self-Control	.79
Sense of Responsibility	.79
Sense of Service	.70
Sociability	.90
Tenacity	.72

6.2. Test-retest Reliability

Test-retest reliability involves administering a test twice to the same respondents, under similar conditions, and comparing the results for each individual in order to determine the stability of the test scores over time. The more similar the scores on the first administration (time 1) are to the scores on the second administration (time 2), the greater the reliability. The Pearson Product Correlation Coefficient gives us the correlation between the two administrations, and ranges between 0 and 1. The closer the correlation is to 1, the greater the reliability.

The formula for the Pearson Product Moment Correlation is:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

where $\bar{x}$ and $\bar{y}$ are the two samples' means, and X and Y are the values of each of the two variables.

For **WORK PROFILE**, test-retest reliability was calculated on an English-speaking sample (n=124) twice, with an interval of 2-3 weeks. Participants ranged in age from 19 to 72 years (mean = 35.01 , sd = 9.67). Test retest reliability was computed from individual scores for all factors, and calculated using the Pearson Moment Correlation Coefficient. Table 11 shows the correlation coefficients between time 1 and time 2.

Table 11. Test-retest Reliability of All Factors.

Factor	Reliability
Adaptability	.66
Ambition	.60
Assertiveness	.65
Autonomy	.56
Compliance	.65
Dynamism	.59
Intellectual Curiosity	.65
Leadership	.78
Organisation	.62
Persuasiveness	.63
Self-Control	.58
Sense of Responsibility	.61
Sense of Service	.73
Sociability	.69
Tenacity	.45

The test-retest correlations for **WORK PROFILE** show good results, with the lowest correlation being at .45, the highest at .78, and the average at .63. It can therefore be concluded, that **WORK PROFILE** is a reliable measure over time, as a person taking the test at two different points in time, will consistently get the same results.

7. Validity

Validity refers to the extent to which an assessment measures what it claims to measure, and that this is well-founded and accurately corresponds to actual behaviour. In psychometrics, validity relates to the degree to which practical evidence and scientific theory support the interpretations of test scores. Whilst the reliability of an assessment relates to the degree to which results are consistent over time, validity tells us the degree to which results of an assessment are accurate and enable us to confidently make predictions about a target outcome. Therefore, a test can be reliable but not valid (one can consistently get the wrong answer) but it cannot be valid without being reliable, as confidence in the accuracy of results depends on the consistency with which we obtain the same results. Established validity protects practitioners from faulty, costly, unethical and legally risky decisions.

Validity is derived from the Latin word ‘validus’, meaning strong. According to the Standards for Educational and Psychological Testing (Standards; APA, AERA, & NCME, 2014), validation is a process of providing evidence for the scientific basis of interpreting assessment scores. The process of validation therefore depends on evidence supporting: a) test content, b) response processes, c) internal structure, d) relations to other variables, and e) consequences of testing. **WORK PROFILE** was validated by following a thorough multistage process that ensured items and factors were accurate, representative, consistent, clear, fair and culturally-free.

Face validity is a measure of whether an assessment looks like it is measuring what it claims to be measuring, by those who take it (Nevo, 1985). This is an important aspect to take into account in order to obtain the candidate’s co-operation and commitment. Face validity and attitudes towards tests influence candidates’ motivation, concentration and test-taking anxiety (McCarthy & Goffin, 2003). As a result, the question of ‘Does **WORK PROFILE** look, to potential candidates, like it measures an individual’s personality profile as well as professional aspirations?’ was addressed during the creation of the assessment, by way of ensuring that aspects such as the situational questions were not misleading or confusing to the user. The following sections will present, in more detail, further aspects of validity that were taken into account for **WORK PROFILE**.

7.1. Concurrent Validity

Concurrent validity measures whether a method is concurrent with the results of another similar, tried and tested method. In this case, in order to establish concurrent validity for **WORK PROFILE**, the components of the assessment must correlate highly with corresponding components on other established personality-based questionnaires. Concurrent validity also requires that the two tests are taken at the same time, or with little time between them, so that the correlation is pure and unaltered by changes over time. Table 12 details the results of the concurrent validity study of **WORK PROFILE** and the Mini-IPIP five-factor personality scale (Goldberg, 1999).

7.1.1. The Mini-IPIP Five-factor Personality Scale

The International Personality Item Pool (IPIP) is a collection of over 3,000 internationally recognised personality items. The project was initiated in 1996 and is available in over 25 languages. The IPIP questionnaire for this study was extracted from the IPIP website (Goldberg, 1999).

The Mini-IPIP five-factor personality scale was designed to be a short version measurement of the Five Factor Model (McCrae & Costa, 1996). The Mini-IPIP consists of 20 items, measuring the five factors of the FFM; *Extraversion, Agreeableness, Conscientiousness, Neuroticism and Imagination*.

Table 12. Factor Correlations between Work Profile and the Mini IPIP Scale.

Factor	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Imagination
Adaptability	0.35	0.12	0.04	-0.28	0.10
Ambition	0.13	-0.01	0.13	-0.09	0.27
Assertiveness	0.10	0.22	0.11	-0.08	0.18
Autonomy	-0.38	-0.22	0.08	0.06	-0.07
Compliance	-0.03	0.04	0.06	0.07	-0.13
Dynamism	0.15	-0.15	-0.10	0.07	-0.01
Intellectual Curiosity	0.16	0.12	-0.12	-0.01	0.16
Leadership	0.45	-0.18	-0.05	-0.12	0.09
Organisation	-0.19	-0.02	0.41	-0.26	0.19
Persuasiveness	0.10	-0.08	-0.38	0.24	-0.06
Self-Control	-0.09	-0.19	0.19	-0.44	0.03
Sense of Responsibility	0.02	-0.03	-0.25	0.18	-0.03
Sense of Service	-0.09	0.36	-0.12	0.08	-0.18
Sociability	0.39	0.32	-0.18	-0.12	-0.06
Tenacity	-0.06	0.03	0.35	-0.33	0.23

The correlations shown in Table 12, between the factors of **WORK PROFILE** and the Mini-IPIP scale, show very good results; ranging from .18 to .45. The strongest correlations lie between the factors *Leadership* and *Extraversion* (.45), which could be expected owing to leaders often being extroverted, and negatively between *Self-Control* and *Neuroticism* (-.44), which is also in line with previous research that finds that neurotic behaviours are linked largely with a lack of self-control. The third highest correlation is found between *Organisation* and *Conscientiousness* (.41); it is not surprising that someone who scores highly on conscientiousness would have a tendency to work in an organised manner. A strong negative correlation was also found between *Persuasiveness* and *Conscientiousness* (-.38). This was unexpected, but could be a result of individuals scoring highly on conscientiousness (i.e. those who like order in their lives), also preferring to verbalise their thoughts and ideas in a straightforward and ordered manner, as opposed to being persuasive and decorating the truth.

7.2. Construct Validity

Construct validity measures the degree to which a test measures what it claims to measure. It refers to the appropriateness of inferences made on the basis of observation or measurement. Construct validity is essential to the overall validity (perceived and actual) of a scale and is particularly valuable when there is much subjectivity to concepts. When a construct is not clearly or thoroughly defined,

we are unable to make direct inferences between the construct and other behaviours (Cronbach & Meehl, 1955). Often there is no accepted unit of measurement for constructs; even fairly well-known concepts such as intelligence are open to debate. Therefore, the construct validity of a scale may be threatened if the construct is: a) defined as too broad or too narrow, b) mislabelled as a specific dimension of or component within the construct, c) confounded by other related constructs that mask the effects of the main construct, or d) the construct was developed with insufficient research. Therefore, in order to develop a scale that measures occupational personality styles, the construct must reflect known theories that define personality, identify the components that it is comprised of and pinpoints positive and negative attributes associated with each dimension.

For work-based assessments, construct validity is particularly important as it: a) adds legitimacy and acceptability to the standardised measurement of employee behaviours, b) facilitates understanding of organisational implications of findings, and c) enables organisations and individuals to tailor interventions to meet desired performance targets. Construct validity addresses the question 'Does **WORK PROFILE** define and classify work-based personality and motivation in a way that represents theoretical and practical understanding?' For **WORK PROFILE**, construct validity was established using two approaches: factor correlations and confirmatory factor analysis.

7.2.1. Inter-factor Correlations

The factors in **WORK PROFILE** measure workplace personality traits, however, they all measure different factors of personality. In other words, if someone scores highly on one factor, it does not mean that they will score highly on all factors. The correlation coefficient is a numerical way to quantify the relationship between two variables, and ranges between -1 and 1. The closer the value to 1, the stronger the relationship. The correlation coefficient does not imply causality; it merely signifies the association between two variables. Traditionally, factor correlations are measured using Pearson (1895) correlation.

Table 13. Inter-factor Correlations.

	Adaptability	Ambition	Assertiveness	Autonomy	Compliance	Dynamism	Intellectual Curiosity	Leadership	Organisation	Persuasiveness	Self-control	Sense of Responsibility	Sense of Service	Sociability	Tenacity
Adaptability															
Ambition	0.09														
Assertiveness	-0.19	0.03													
Autonomy	-0.17	-0.24	0.01												
Compliance	-0.14	-0.10	-0.24	-0.04											
Dynamism	0.17	0.15	0.11	0.22	-0.33										
Intellectual Curiosity	0.18	-0.06	0.01	-0.08	-0.11	0.04									
Leadership	0.11	0.22	0.24	-0.25	-0.29	0.11	-0.16								
Organisation	-0.25	0.03	-0.03	0.07	0.17	-0.28	-0.03	-0.04							
Persuasiveness	-0.21	0.08	-0.24	-0.05	-0.07	-0.04	-0.03	0.08	-0.06						
Self-Control	-0.01	-0.02	-0.33	0.01	0.24	-0.16	0.01	-0.06	0.28	-0.14					
Sense of Responsibility	0.02	0.11	-0.01	-0.08	-0.27	0.01	0.01	0.18	-0.30	0.10	-0.11				
Sense of Service	0.03	-0.37	-0.06	-0.34	0.17	-0.24	-0.09	-0.35	-0.12	-0.05	0.00	0.04			
Sociability	0.30	0.06	0.13	-0.37	-0.28	-0.11	0.15	0.10	-0.29	-0.01	0.19	-0.10	0.28		
Tenacity	0.08	0.12	-0.14	-0.09	0.06	0.00	-0.03	0.08	0.09	-0.21	0.02	-0.02	0.03	0.06	

7.2.2. Confirmatory Factor Analysis

To investigate the theoretical constructs or factors that may be represented by a set of items as well as the quality of individual items, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are usually performed. EFA and CFA are both forms of factor analysis. Factor analysis examines the factors (hypothetical constructs) of an assessment which are developed to interpret clusters of strong inter-correlations between items in a correlation matrix. Factor analysis aims to simplify a range of items into a few common variables; these can be identified by extracting and grouping items which have high correlations on a factor to reveal the common variables.

The goal of EFA is to identify and maximise the importance of each dimension (percentage of variance), while CFA is a hypothesis-driven approach that determines the degree to which observed indicators reflect the underlying construct by specifying the number of factors expected and then classifying the value of expected relationships (Schreiber et al, 2006). In doing so, CFA provides a numerical estimate of the degree to which the proposed factor structure is an accurate fit of the phenomenon in the real world. CFA therefore investigates the degree to which measures of a construct are consistent with established understanding and theoretical arguments for that construct.

A confirmatory factor analysis (CFA) with continuous factor indicators was conducted in order to check the convergence of the results with the model of the assessment. CFA is a form of factor analysis used to measure whether a construct is consistent with its intended meaning. There are two key approaches to confirming the factor structure of a scale. A strict confirmatory approach is one in which a model is tested using structural equation modelling (SEM) goodness of fit tests to determine if the model specified by the research is consistent with the data. On the other hand, an alternative model approach allows two or more models to be tested, in order to identify which has the best fit. For the development of **WORK PROFILE**, a strict confirmatory approach was adopted as there is only one formulated model for **WORK PROFILE**.

CFA is used to determine how well the model fits the data, and therefore evaluating the goodness of fit indices is crucial. Several statistical tests are used to determine how well the model fits the data in CFA and SEM. The closer fit indices meet statistical requirements, the better the model. The Chi Square (X^2) statistic is considered the primary index of goodness-of-fit. However, due to problems associated with reliance on the X^2 (sensitivity to sample sizes and inappropriateness when distributional assumptions are violated), measures of relative fit are strongly recommended. The X^2 is a measure of absolute fit, and provides a fundamental indication of how well the proposed theory fits the data and does not rely on comparison with a baseline model (Jöreskog & Sörbom, 1993).

For the validation of **WORK PROFILE**, the Root Mean Square Error of Approximation (RMSEA, Steiger & Lind, 1980), the Bentler Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI) were chosen as additional measures of fit, as they are the least likely to be affected by sample size (Fan, Thompson & Wang, 1999). Furthermore, using a variety of indices is necessary as different indices reflect different aspects of model fit (Hooper, Coughlan & Mullen, 2008). Like the X^2 , the RMSEA is a measure of absolute fit that assesses the amount of error associated with a proposed model. The lower the RMSEA, the more robust the model; indices of .05 or below are considered to be a good fit (Hu & Bentler, 1999). This is because RMSEA is related to the residuals in the model and assesses the degree to which the proposed model is inaccurate. Both the CFI and TLI are incremental or relative fit indices that assess overall improvement of a proposed model compared with an independence model where observed variables are uncorrelated (Hooper et al., 2008; Holtzman & Vezzu, 2011). While higher values of the CFI and TLI (.9 and above) indicate an excellent fit, it should be noted that a CFI or TLI of .90 or higher may not necessarily correspond to a RMSEA of .05 or less (Hu & Bentler, 1999).

Due to the Ipsative nature of the test, the method of Thurstonian IRT (described in chapter 2.2.4.) was used to conduct the CFA.

Table 14. The Goodness of Fit Indices and Model Results.

N	X ²	Df	X ² /df	P	RMSEA	P<=.05	CFI	TLI
373	4202.233	3659	1.148	0	.02	1	.887	.877

The results of the Thurstonian IRT, displayed in Table 14, shows an RMSEA of of .02 and CFI of .887, indicating *that* **WORK PROFILE** has a good model fit and that the factors exist in the defined structure.

8. WORK PROFILE and Equal Opportunities

Organisations are increasingly under pressure to ensure that tools used to inform employee recruitment, selection, development, promotion and performance measurement decisions are fair, objective and accurate. This pressure stems from a progressively knowledgeable applicant pool, a more sophisticated labour force and stringent local and international employment legislation and testing standards. Equal opportunities therefore represent a critical consideration for any modern day work assessment, and an assessment is only considered to be truly valid when it is appropriate for use with different groups. **WORK PROFILE** is a fair and accurate measure of personality, and HR practitioners and psychologists can be confident of its diverse application.

Within the context of assessment, measures used must not result in adverse impact when applied to different groups. Adverse impact occurs when there is a significant difference in the rate of selection and employment decisions for one protected group over another. This is particularly important for organisations as there is a risk of selecting the wrong candidate for a job, as well as legal and financial costs associated with discrimination claims. Adverse impact is classified as discrimination when a candidate believes (and validity evidence fails to show otherwise) that not being offered a role was based on membership to certain demographics. For most countries across Europe, North America and Australia, fairness in the selection process is subject to legislation. Specifically for the USA, the Employee Testing and other Procedures Act 1971 (US Department of Labor) protects candidates against unfair assessment procedures, whilst fair employment assessment in the UK is protected by the following:

- Race Relations Act (1976)
- Sex Discrimination Act (1975;1986)
- Disability Act (1998)
- Employment Equality Regulation (2003)
 - Sexual Orientation
 - Religion and Belief
- Age Discrimination (2006)

8.1. Measurement Equivalence

Measurement equivalence (ME) refers to the consistency of measurement across some specified group demarcation such as nationality, ethnicity and gender. The focus of ME is to ensure that individuals from different groups with the same standing on an attribute have the same set of observations on an assessment (Drasgow, 1987; McLeary & Cruise, 2012). Understanding ME is important in assessments in order to establish validity, confidence and equivalence across groups. Equivalence of measurement is obtained when candidates from different cultural backgrounds, age groups, ethnicities, genders and occupational groups with the same level on a trait, have an equal chance of being selected for a role that requires the desired level on this trait. ME helps to determine the extent to which respondents are interpreting items similarly, and identified differences across groups reflect actual differences in the trait being measured (Drasgow, 1984). One approach to assessing measurement equivalence is differential item functioning analysis.

8.1.1. Differential Item Functioning (DIF)

An item functions differently, or displays DIF, if respondents from different groups have different probabilities of endorsing the item, despite having the same level of the attribute or trait. When a different probability of endorsing an item remains constant across different levels of the attribute or

trait, the DIF is said to be uniform. Here, the item of interest consistently gives one group an advantage across all levels of the attribute or trait. Conversely, when a different probability of endorsing an item varies along the level of the attribute or trait, the DIF is characterised as non-uniform (Clauser & Mazor, 1998; Hanson, 1998). Rather than a consistent advantage being given to the reference group across the ability continuum, an advantage to one group over another exists at different levels of the trait. For instance, an item may give the reference group an advantage at the higher end of the continuum, while a disadvantage at the lower end (see Figure 6). Item Response Theory (IRT) facilitates DIF analysis by examining the probability of endorsing items through the item characteristic curve. Here, the logistic regression approach is used, whereby separate analyses are run for each item and maximum likelihood estimation used to calculate the probability of endorsing an item. Figures 5 and 6 show an example of item characteristic curves for uniform and non-uniform DIF.

8.1.1.a. DIF Analysis: Gender

DIF analysis was performed across gender to ensure that there was not a bias towards either males or females, and therefore that **WORK PROFILE** is a fair test for both genders. For this analysis, the norm sample was divided according to gender; 196 females and 177 males.

Table 15. Coefficients for Significant DIF and Effect Size for DIF Items.

Factor	Coefficient	R ²	Effect size
Ambition	7.46	0.03	Negligible
Assertiveness	6.71	0.02	Negligible
Autonomy	12.89	0.04	Negligible
Autonomy 2	16.32	0.04	Negligible
Compliance	6.00	0.02	Negligible
Intellectual Curiosity	13.84	0.04	Negligible
Intellectual Curiosity 2	14.57	0.05	Negligible
Leadership	7.37	0.02	Negligible
Sense of Service	7.65	0.02	Negligible
Sense of Service 2	7.44	0.02	Negligible
Tenacity	7.90	0.02	Negligible
Tenacity 2	14.67	0.04	Negligible
Tenacity 3	8.06	0.02	Negligible

The analysis flagged 13 of the 90 items for DIF, based on gender. Table 15 shows the DIF coefficient as well as the R² effect size. According to Zumbo and Thomas' (1997) classification of effect size, any R² under .13 is considered negligible. Out of the 13 items flagged for DIF, none showed an effect size of .13 or over. Therefore, it can be considered that **WORK PROFILE** does not contain any item that functions differently across genders. Two items (figure 6 and Figure 7) have been chosen to

demonstrate uniform and non-uniform DIF. Figures 5 and 6 show uniform and non-uniform DIF, respectively.

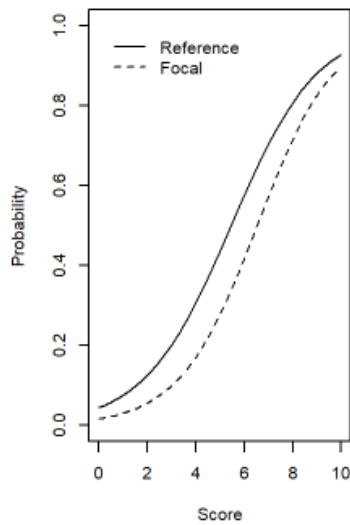


Figure 5. Uniform DIF for the Tenacity

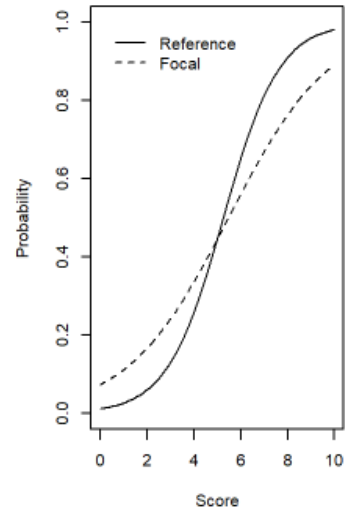


Figure 6. Non-Uniform DIF for Sense of Service

8.1.1.b. T-tests and ANOVAS

The DIF analysis demonstrates that **WORK PROFILE** has high consequential validity, and although several items were flagged with DIF, their impact on the overall assessment is negligible, meaning that the assessment does not bias towards gender. Therefore, any significant differences found between genders can be attributed to real differences between the groups, rather than biases of the assessment. To further explore the differences found between the groups, an independent samples t-test was performed between the groups, t-tests and ANOVAS were performed between genders, age groups and study levels (see Tables 16 to 18).

Gender

In order to check for differences in scores between genders, an independent samples t-test was conducted to compare all **WORK PROFILE** factors between females (n = 196) and males (n = 177) (see Table 16).

Table 16. T-tests for Factors by Gender.

Factor	Female	Male	T.	df	p
Adaptability	4.78	4.79	-.03	371	.97
Ambition	5.28	5.74	-1.93	371	.05
Assertiveness	5.67	5.41	.95	371	.34
Autonomy	5.83	5.62	.87	371	.39
Compliance	5.51	5.39	.47	371	.64

Dynamism	5.21	5.00	.86	371	.39
Intellectual Curiosity	6.20	6.72	-2.10	371	.04*
Leadership	3.43	4.33	-3.08	371	.00***
Organisation	5.70	5.81	-.44	371	.66
Persuasiveness	3.78	4.14	-1.51	371	.13
Self-Control	5.63	5.94	-1.19	371	.24
Sense of Responsibility	4.39	4.04	1.37	371	.17
Sense of Service	4.98	3.90	3.79	371	.00***
Sociability	5.3	5.03	.83	371	.41
Tenacity	5.23	5.32	-.44	371	.66

Note: *p < .05, ** p < .01, ***p < .001.

The t-tests found significant differences between males and females on the factors *Intellectual Curiosity*, *Leadership*, and *Sense of Service*. Males were found to score significantly higher on *Intellectual Curiosity* and *Leadership*, but lower on *Sense of Service*.

Study level

Participants were separated into five groups according to their study level: Masters and PhDs (n=87), Bachelor's degrees (n=172), 2 years vocational study (n=67), High school diploma (n=42) and trades training without study (n=2), students (n=3).

Table 17. ANOVA for Factors by Study Level.

Factor	Master's Degree / PhD	Bachelor's degree	2 years vocational study	High school diploma	Trades Training / No studies	Student	F	df	p-value
Adaptability	4.89	4.81	4.71	4.53	6.15	4.37	0.37	5	0.87
Ambition	5.9	5.54	5.27	4.86	4.6	5.57	1.37	5	0.24
Assertiveness	5.94	5.27	5.74	5.52	7.1	4.73	0.97	5	0.43
Autonomy	5.33	5.76	5.76	6.39	7.5	5.3	1.39	5	0.23
Compliance	4.56	5.68	5.55	6.15	3.75	6.97	3.81	5	0
Dynamism	4.87	5.15	5.09	5.52	4.55	4.47	0.49	5	0.78
Intellectual Curiosity	6.73	6.38	6.06	6.85	5.85	5.53	0.98	5	0.43
Leadership	4.62	3.94	3.21	3.04	2.95	3.6	2.8	5	0.02

Organisation	5.69	6.1	5.38	5.22	4.2	4.47	1.59	5	0.16
Persuasiveness	4.35	4.18	3.44	2.85	4.6	5.57	3.82	5	0
Self-Control	5.33	5.95	5.49	6.48	2.9	7.2	2.41	5	0.04
Sense of Responsibility	4.21	4.17	4.44	4.11	3.75	4.43	0.16	5	0.98
Sense of Service	4.04	4.48	4.94	4.34	4.15	7.77	1.68	5	0.14
Sociability	6.29	4.88	5.17	4.16	2.9	5.3	3.79	5	0
Tenacity	4.88	5.33	5.38	5.7	5.4	5.53	1.05	5	0.39

For an in-depth analysis of the differences between study groups for the **WORK PROFILE** factors, an ANOVA was conducted for all factors (Table 17). The ANOVA found that there were significant differences for five factors based on study level, which were *Compliance*, *Leadership*, *Persuasiveness*, *Self-Control* and *Sociability*. Post hoc analysis was conducted using pair-wise t-tests (bonferroni correction) in order to see where these differences occurred.

For the factor *Compliance*, the Master's level group (M=4.56) was significantly lower than the Bachelor's (M=5.68, $p=.007$) and High school level groups (M=6.15, $p=.008$). All other comparisons showed no significant differences.

For *Leadership*, the Master's level (M=4.62) was significantly higher than the Vocational (M=3.21, $p=.032$) and High school level groups (M=3.04, $p=.043$). All other comparisons showed no significant differences.

For *Persuasiveness*, the High school level (M=2.85) was significantly lower than the Master's (M=4.35, $p=.008$) and Bachelor's levels (M=4.18, $p=.012$). All other comparisons showed no significant differences.

For *Self-Control*, the Bonferroni corrected post hoc t-tests showed there were no significant differences.

For *Sociability*, the Master's level (M=6.29) was significantly higher than the Bachelor's (M = 4.88, $p=.008$) and High school Levels (M=4.16, $p=.004$). All other comparisons showed no significant differences.

Age

Participants were separated into six groups according to their age: 18-25 (n=52), 26-33 (n=141), 34-41 (n=74), 42-49 (n=42), 50-56 (n=28) and 57+ (n=21).

Table 18. ANOVA Table for Age.

Factor	18-25	26-33	34-41	42-49	50-56	57+	NA	F	df	p-value
Adaptability	4.63	4.74	4.33	4.90	5.11	5.50	6.01	1.95	6.00	0.07

Ambition	5.76	5.83	5.46	4.95	4.41	4.88	6.11	2.48	6.00	0.02
Assertiveness	5.07	5.46	5.85	5.72	6.31	6.14	3.73	2.19	6.00	0.04
Autonomy	5.40	5.42	6.06	6.47	6.14	6.15	4.81	2.10	6.00	0.05
Compliance	5.20	5.38	5.52	5.70	5.84	4.76	6.22	0.81	6.00	0.56
Dynamism	5.25	5.13	4.96	4.86	5.21	5.36	5.27	0.21	6.00	0.97
Intellectual Curiosity	7.15	6.24	6.56	6.23	6.58	5.87	6.55	1.25	6.00	0.28
Leadership	3.40	4.12	3.75	3.31	4.05	4.05	4.39	0.83	6.00	0.55
Organisation	5.17	5.85	6.02	6.22	5.65	5.20	5.18	1.14	6.00	0.34
Persuasiveness	4.20	4.06	4.14	4.05	2.71	3.65	3.61	1.71	6.00	0.12
Self-Control	5.81	5.90	5.57	6.28	5.15	5.39	5.85	0.84	6.00	0.54
Sense of Responsibility	4.27	4.48	3.85	4.17	3.99	4.01	4.39	0.61	6.00	0.72
Sense of Service	4.42	4.55	4.14	4.45	4.72	5.24	3.99	0.57	6.00	0.76
Sociability	5.23	5.12	5.22	4.68	5.09	6.30	5.18	0.65	6.00	0.69
Tenacity	5.41	5.05	5.15	5.00	6.37	5.52	5.87	2.03	6.00	0.06

Note: NA = Undefined age

The ANOVA found that there were significant differences for the factors *Ambition* and *Assertiveness*. Post hoc analysis was conducted using pair-wise t-tests (Bonferroni correction) in order to see where these differences occurred. However, these showed that there were no significant differences between the groups.

9. Additional Information and Considerations

WORK PROFILE was developed primarily to guide Human Resource decisions for recruitment but its usage is manifold for organisational settings. It is important that the appropriate norms, scores and interpretations are used in order to aid selection, career development, coaching, and performance management. Accurate interpretation of psychometric tests helps both administrators and candidates to: a) put the assessment into context; b) effectively compare candidates across target capabilities; c) accurately identify areas for development; d) maximise fit between job requirements and candidate tendencies and aspirations; and e) promote the organisation as a fair and objective place to work. The current chapter gives guidelines for Human Resource practitioners to understand and successfully interpret candidate scores of **WORK PROFILE** for different organisational objectives.

Report generation software for **WORK PROFILE** provides graphs and personalised comments for all the 15 factors that relate to the candidate's score. The report should only be interpreted by professionals who have completed *Central Test*'s accredited training course and are familiar with this manual. Training in and knowledge of the assessment process are vital as scores should be interpreted in light of information regarding reliability, validity, standard error of measurement, and any accommodations to the assessment or assessment session that were made. Scores should also be presented within the context of the assessment and in a clear, meaningful and appropriate way for the target audience.

9.1. Interpreting Scores

It is important to note that all the assessments by *Central Test* have been developed for Human Resource purposes only. Therefore, no clinical conclusions can be drawn from the results – not even from extreme scores. There are six different levels of comments in total and these levels are determined by a candidate's score. Table 19 provides an example of the descriptors used to describe different levels of a factor. Extreme scores are defined as the lowest and highest possible values on a factor. For **WORK PROFILE**, the minimum score is set at 0 and the maximum score is set at 10. Given the ipsative format of the assessment, it is impossible to obtain extreme scores on all factors.

High scores on a factor indicate that there is a high probability that the person who is being evaluated can be characterised as having a tendency/preference for that factor. Middle scores (i.e. 4-7) indicate a high congruence with the norm group and may be less of a tendency/preference for the individual. However, all comments should be viewed as hypotheses that need to be confirmed with other sources of data, such as interviews or further assessments, CV content and references.

Table 19. Example Descriptors of Different Levels of a Trait or Inclination in Report Comments for Work Profile.

Level 1 (Scores 0 to 2)	Level 2 (Scores 2.1 to 3.5)	Level 3 (Scores 3.6 to 5.0)	Level 4 (Scores 5.1 to 6.4)	Level 5 (Scores 6.5 to 7.8)	Level 6 (Scores 7.9 to 10)
Does not seem to indicate a tendency	Indicates a very low tendency	Indicates a low tendency, may occasionally enjoy	Indicates a tendency, but not a preference	Indicates a tendency	Indicates a strong tendency

9.2. Organisational Need: Career Selection, Recruitment and Development

Through the identification of 21st century organisational requirements, **WORK PROFILE** is an effective tool for both the employer, in the selection of job candidates on the basis of a personality profile as well as professional aspirations, and for the job candidate, in the successful selection of an appropriate career navigation.

In order to successfully hire a candidate that will feel satisfied with their job and as a result want to remain with the company, job satisfaction of the employee becomes significant to the recruiter. By using **WORK PROFILE**, an employer will be able to infer how appropriate a job may be for a candidate based on an objective assessment of their tendencies and preferences. If the endorsed attributes of said candidate do not align with those of the job role in question, there may not be a good match between the two.

Recruitment and selection is an important organisational function which can be very time-consuming and expensive. Furthermore, bad recruitment decisions have additional effects on organisational productivity, capability and sustainability. Whilst someone in financial and legal administration, for example, may feel more comfortable if they have a natural orientation for tasks that involve following strict rules; someone working in research and development may be better suited to have a preference in tasks that involve inquisitiveness and curiosity. It is recommended that prior to administering **WORK PROFILE**, recruiters have a detailed job description and person specifications based on a job analysis.

WORK PROFILE helps recruiters to customise job profiles in their client account by specifying desired scores for each factor. The assessment allows the recruiter to easily compare the target profile with a candidate's profile.

Furthermore, **WORK PROFILE** can be used in line with performance evaluations, which are a key component of employee development and success. Performance reviews should be fair and balanced assessments of an employee's achievements and capabilities, and should give the opportunity to discuss expectations and accomplishments. **WORK PROFILE** supports Human Resource practitioners with performance reviews that can help to discuss job performance in relation to personality profiles and professional aspirations. If an individual is not performing as they could be at work, it may be down to a lack of a preference in the particular area that they are working on. Identifying this can help the recruiter to identify whether or not the employee should be transferred to another area of the company, or to a different aspect of a project.

Whilst **WORK PROFILE** is effective in identifying areas for development, it should not be used to directly track an employee's performance progress. We therefore recommend the use of objective performance assessments or other evaluations in conjunction with **WORK PROFILE**.

References

- Allport, G. W. (1937). *Personality: a psychological interpretation*. H. Holt and Company.
- Alonso, P., & Lewis, G. B. (2001). Public Service Motivation and Job Performance Evidence from the Federal Sector. *The American Review of Public Administration*, 31(4), 363-380.
<http://doi.org/10.1177/02750740122064992>
- Anderson, E. W., Fornell, C., & Mazvancheryl, S. K. (2004). Customer Satisfaction and Shareholder Value. *Journal of Marketing*, 68(4), 172-185. <http://doi.org/10.1509/jmkg.68.4.172.42723>
- Barnard, C. I. (1968). *The functions of the executive*. Cambridge: Harvard University Press.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology*, 44(1), 1-26. <http://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Bowen, C., Martin, B. A., & Hunt, S. T. (2002). A comparison of ipsative and normative approaches for ability to control faking in personality questionnaires. *The International Journal of Organizational Analysis*, 10(3), 240-259. <http://doi.org/10.1108/eb028952>
- Boyatzis, R. E., & Skelly, F. R. (1991). The impact of changing values on organizational life. In A. D. Kolb, I. M. Rubin, and J. S. Osland (Eds.). *The Organizational Behavior Reader*. Englewood Cliffs, NJ: Prentice Hall Press
- Bright, L. (2007). Does Person-Organization Fit Mediate the Relationship Between Public Service Motivation and the Job Performance of Public Employees? *Review of Public Personnel Administration*, 27(4), 361-379. <http://doi.org/10.1177/0734371X07307149>
- Brown, A. (2010). *How Item Response Theory can solve problems of ipsative data (Doctoral dissertation)*. Universitat de Barcelona, Barcelona.
http://www.tdx.cat/bitstream/handle/10803/80006/ANNA_BROWN_PhD_THESIS.pdf;jsessionid=CC2F4AC92124CAFDFE575BA243C3938D.tdx1?sequence=1
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, 71(3), 460-502.
<http://doi.org/10.1177/0013164410375112>
- Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods*, 18(1), 36-52.
<http://doi.org/10.1037/a0030641>
- Cattell, R. B. (1957). *Personality and Motivation Structure and Measurement*. New York: World Book.
- Chan, W. (2003). Analyzing ipsative data in psychological research. *Behaviormetrika*, 30(1), 99-121.
<http://doi.org/10.2333/bhmk.30.99>
- Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *The Academy of Management Review*, 14(3), 333-349.
<http://doi.org/10.2307/258171>

- Chiu, C. H. (2002). Relationship between *Employee Attitudes and Alignment of Employees-Organizational Values (Doctoral dissertation)*, Graduate Institute of Healthcare Organization Administration, National Taiwan University.
- Chung, Y. I., Chi, N. W., Chen, C. Y. (2008). Development and Valuation of Measurement on Work Values of Sales. *NTU Management Review*, 19(1): 51-82.
- Clauser, B. E., & Mazor, K. M. (1998). Using Statistical Procedures to Identify Differentially Functioning Test Items. *Educational Measurement: Issues and Practice*, 17(1), 31-44.
<http://doi.org/10.1111/j.1745-3992.1998.tb00619.x>
- Cornwell, J. M., & Dunlap, W. P. (1994). On the questionable soundness of factoring ipsative data: A response to Saville & Willson (1991). *Journal of Occupational and Organizational Psychology*, 67(2), 89-100.
<http://doi.org/10.1111/j.2044-8325.1994.tb00553.x>
- Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa., P. T., Jr., & McCrae. R. R. (1987). Personality assessment in psychosomatic medicine: Value of a trait taxonomy. In G. A. Fava & T. N. Wise (Eds.). *Advances in psychosomatic medicine: Vol. 17. Research paradigms in psychosomatic medicine*. (pp. 71-82). Basel, Switzerland: Karger.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
<http://doi.org/10.1037/h0040957>
- Cruise, P. A. (2011). The Role of Culture in Personnel Selection: Does culture affect hiring the right person for the job? *Lambert Academic Publishing*. Germany.
- Cruise, P. A. (2012). The stability and organisational value of personality assessment. *Psychometrics Forum UK*.
- De Fruyt, F., & Mervielde, I. (1999). RIASEC types and big five traits as predictors of employment status and nature of employment. *Personnel Psychology*, 52(3), 701-727.
<http://doi.org/10.1111/j.1744-6570.1999.tb00177.x>
- Dodou, D., & de Winter, J. C. F. (2014). Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior*, 36, 487-495.
<http://doi.org/10.1016/j.chb.2014.04.005>
- Drasgow, F. (1987). Study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology*, 72(1), 19-29.
<http://doi.org/10.1037/0021-9010.72.1.19>
- Edwards, A. L., & Kenney, K. C. (1946). A comparison of the Thurstone and Likert techniques of attitude scale construction. *Journal of Applied Psychology*, 30(1), 72-83.
<http://doi.org/10.1037/h0062418>

- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 56-83.
<http://doi.org/10.1080/10705519909540119>
- Finegan, J. E. (2000). The impact of person and organizational values on organizational commitment. *Journal of Occupational and Organizational Psychology*, 73(2), 149-169.
<http://doi.org/10.1348/096317900166958>
- Galton, S. F. (1884). The Measurement of Character. *The Fortnightly Review*.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.
- Griffin, Em. (1991). *A First Look at Communication Theory*. McGraw-Hill, Inc.
www.afirstlook.com
- Harvill, L. M. (1991). Standard Error of Measurement. *Educational Measurement: Issues and Practice*, 10(2), 33-41.
<http://doi.org/10.1111/j.1745-3992.1991.tb00195.x>
- Holland, J. L. (1959). A theory of vocational choice. *Journal of counseling psychology*, 6(1), 35.
- Holtzman, S., & Vezzu, S. (2011). Confirmatory Factor Analysis and Structural Equation Modeling of Noncognitive Assessments using PROC CALIS. Retrieved from <http://www.lexjansen.com/nesug/nesug11/sa/sa07.pdf>
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods*, 6, 53-60.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: the Big Five revisited. *The Journal of Applied Psychology*, 85(6), 869-879.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research* (L. Pervin and O.P. John). Elsevier.
- Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. *Scientific Software International*.
- Kaplan, R. M., & Saccuzzo, D. P. (2008). *Psychological Testing: Principles, Applications, and Issues* (7th ed.). Belmont, CA: Wadsworth Publishing Co Inc.
- Kluger, A. N., Reilly, R. R., & Russell, C. J. (1991). Faking biodata tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology*, 76(6), 889-896.
<http://doi.org/10.1037/00219010.76.6.889>

- Kovach, K. A. (1995). Employee motivation: Addressing a crucial factor in your organization's performance. *Employment Relations Today*, 22(2), 93-107.
<http://doi.org/10.1002/ert.3910220209>
- Lee, K., & Ashton, M. C. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, 39(2), 329-358.
http://doi.org/10.1207/s15327906mbr3902_8
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705-717.
<http://doi.org/10.1037/0003066X.57.9.705>
- Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin*, 90(1), 125-152.
<http://doi.org/10.1037/0033-2909.90.1.125>
- Loo, R. (1999). Issues in factor-analyzing ipsative measures: The Learning Style Inventory (LSI-1985) example. *Journal of Business and Psychology*, 14(1), 149-154.
<http://doi.org/10.1023/A:1022918803653>
- Major, D. A., Turner, J. E., & Fletcher, T. D. (2006). Linking proactive personality and the Big Five to motivation to learn and development activity. *The Journal of Applied Psychology*, 91(4), 927-935.
<http://doi.org/10.1037/0021-9010.91.4.927>
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370-396.
<http://doi.org/10.1037/h0054346>
- Massey, M. E. (1979). *The people puzzle: understanding yourself and others*. Reston Pub. Co.
- McCarthy, J. M., & Goffin, R. D. (2003). Is the Test Attitude Survey Psychometrically Sound? *Educational and Psychological Measurement*, 63(3), 446-464.
<http://doi.org/10.1177/0013164403063003007>
- McCrae, R. R., & John, O. P. (1992). An introduction to the five factor model and its applications. *Journal of Personality*, 176-207.
http://psych.colorado.edu/~carey/courses/psyc5112/readings/psnbig5_mccrae03.pdf
- McLeary, C., & Cruise, P. (2012). *Modeling Organisational Trust: Addressing Challenges of Measurement Invariance*. In Proceedings of the 11th European Conference on Research Methods (p. 470). Academic Conferences Limited.
- Meade, A. W. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology*, 77(4), 531-552.
<http://doi.org/10.1348/0963179042596504>
- Mooradian, T., Renzl, B., & Matzler, K. (2006). Who Trusts? Personality, Trust and Knowledge Sharing. *Management Learning*, 37(4), 523-540.
<http://doi.org/10.1177/1350507606073424>

- Mueller-Hanson, R., Heggstad, E. D., & Thornton III, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, 88(2), 348-355.
<http://doi.org/10.1037/0021-9010.88.2.348>
- Nadeau, L., Acier, D., Miranda, D., Pervin, L.-A., & John, O.-P. (2004). *La personnalité : De la théorie à la recherche*. Bruxelles : Saint-Laurent, Québec: De Boeck.
- Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, 22(4), 287-293.
<http://doi.org/10.1111/j.1745-3984.1985.tb01065.x>
- Pardee, Ronald L. (1990). *Motivation Theories of Maslow, Herzberg, McGregor & McClelland: A Literature Review of Selected Theories Dealing With Job Satisfaction and Motivation*. Washington, D.C.: Distributed by ERIC Clearinghouse.
<http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED316767>
- Porter, L. W., Bigley, G. A., & Steers, R. M. (2003). *Motivation and work behavior*. McGraw-Hill.
- Rainall, S. (2004). A review of employee motivation theories and their implications for employee retention within organizations. *The journal of American academy of business*, 9, 21–26.
- Rothmann, S., & Coetzer, E. (2003). The Big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, North America, 29.
- Saville, P., & Willson, E. (1991). The reliability and validity of normative and ipsative approaches in the measurement of personality. *Journal of Occupational Psychology*, 64(3), 219-238.
<http://doi.org/10.1111/j.2044-8325.1991.tb00556.x>
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research*, 99(6), 323-338.
<http://doi.org/10.3200/JOER.99.6.323-338>
- Schneider, B. (1987). The People Make the Place. *Personnel Psychology*, 40(3), 437-453.
<http://doi.org/10.1111/j.1744-6570.1987.tb00609.x>
- Schutz, W. C. (1958). *FIRO: A three-dimensional theory of interpersonal behavior (Vol. ix)*. Oxford, England: Rinehart.
- Schwartz, S. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, 2(1).
<http://doi.org/10.9707/2307-0919.1116>
- Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for the number of common factors*. In annual meeting of the Psychometric Society, Iowa City, IA (Vol. 758).
- Studenska, Anna. (2012). Predictors of Learning Autonomy and its Components. *Literacy Information and Computer Education Journal (LICEJ)*, Volume 3, Issue 1, 599

- Ten Holt, J. C., van Duijn, M. A. J., & Boomsma, A. (2010). Scale construction and evaluation in practice: A review of factor analysis versus item response theory applications. *Psychological Test and Assessment Modeling*, 52(3), 272-297.
- Tett, R. P., & Simonet, D. V. (2011). Faking in Personality Assessment: A "Multisaturation" Perspective on Faking as Performance. *Human Performance*, 24(4), 302-321. <http://doi.org/10.1080/08959285.2011.597472>
- Thompson, Sheila R. (2001). The Five Most Important Entry-Level Employability Skills That Employers of the Chippewa Valley Seek in Entry-Level Job Applicants, 3-6. *American Psychological Association (APA) Publication Manual*.
- Thurstone, L. L. (1927). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, 21(4), 384-400. <http://doi.org/10.1037/h0065439>
- Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review*, 38(5), 406-427. <http://doi.org/10.1037/h0069792>
- Zumbo, B. D., & Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF*. Prince George, Canada: Edgeworth Laboratory for Quantitative Behavioral Science, University of Northern British Columbia.

CONTACT US

Client Service: **customer@centraltest.com**

Training and consulting: **consulting@centraltest.com**



www.centraltest.com

