

TECHNICAL MANUAL

&

USER GUIDE

IMMERSION

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

*The assessment **IMMERSION** and the Central Test logo are registered trademarks.*

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

Table of Contents

TECHNICAL MANUAL	1
Table of Contents	2
Introduction	4
1. General Description and Applicability	6
1.1. Main Areas of Application	10
1.2. Questionnaire Structure	11
1.2.1. Format	11
1.2.2. Questions	11
1.3. Languages and Special Requirements	11
1.3.1. Languages available	11
1.3.2. Accessibility options and special requirements	12
1.4. Social Desirability	12
1.5. Administration	13
1.6. Technical and Security Issues	13
1.6.1. Respecting the Ethical Charter	13
1.6.2. Confidentiality and Security of User Data	13
1.6.3. Hardware Architecture and Server Availability	14
1.6.4. Computer-generated Reports	14
1.7. Supply, Conditions and Availability	15
2. Conceptual Framework and Development	16
2.1. Theoretical Background	16
2.1.2. Links between the RIASEC dimensions and personality traits	16
2.1.3. Subdivision of the RIASEC factors	16
2.2. Gamification of Psychometric Assessments	16
2.2.1. Gamified assessments vs. traditional assessments	16
2.2.2. Management of expectations for gamified assessments	17
2.3. Development of the Tool	17
2.3.1. Factors and Item Construction	17
2.3.2. Combined Profiles	18
2.3.3. Ipsative Response Format	18
2.3.4. Thurstonian Item Response Theory (IRT) and Ipsative Responses	18
2.3.5. Item Difficulty and Discrimination	19
2.3.6. Pretests and Pilot Study	19
3. Scoring	20

3.1. Raw Scores	20
3.2. Conversion of Raw Scores to Standard Scores	20
3.3. Scoring in the Report	20
4. Psychometric Properties	22
4.1. Sample Descriptives	22
4.2. Descriptive Norms	23
4.3. Reliability	29
4.3.1. Internal Reliability	30
4.4. Validity	31
4.4.1. Construct Validity	31
4.4.1.a. Inter-factor Correlations	32
4.4.1.b. Confirmatory Factor Analysis	33
5. IMMERSION and Equal Opportunities	37
5.1. Measurement Equivalence	37
5.1.1. Differential Item Functioning (DIF)	38
5.1.1.a. DIF Analysis: Gender	38
5.1.1.b. DIF Analysis: Age	39
5.1.2. T-tests and ANOVAs	40
5.1.2.a. Age Groups by Factors	40
5.1.2.b. Age Groups by Sub-factors	41
5.1.2.c. Gender	44
6. Additional Information and Considerations	46
6.1. Interpreting Scores	46
6.2. Report Content	47
6.3. Organisational Need: Career Selection, Recruitment and Development	48
References	50

Introduction

IMMERSION is a gamified tool for assessing professional interests, built around scenarios designed to identify individual preferences. Based on John L. Holland's theory of vocational choices (1959), known as the RIASEC model, **IMMERSION** highlights users' core motivations and provides detailed insights into their aspirations and the career fields most likely to match their interests.

The 2025 edition of **IMMERSION** is an interactive update of the earlier non-gamified version previously offered by *Central Test*. Featuring a modern interface, this new edition transforms the traditional assessment format by immersing users in an engaging environment enriched with gamified elements. The goal of this redesign is to offer a smoother, more dynamic, and more engaging assessment experience.



Figure 1. Example of how the gamified assessment is presented from the user's perspective.

Among other features, the visual gamification elements enhance the user experience. This immersive setting stimulates the candidate's attention and motivation, encouraging more spontaneous, accurate and realistic responses compared to a traditional assessment.

Other benefits of **IMMERSION** are listed below:

*Table 1. Description of **IMMERSION** benefits.*

1	Short and easy-to-understand questions reflecting current professional fields.
2	Immersive questions presented through engaging scenarios
3	Quick completion: approximately 15 minutes on average
4	Smooth, intuitive and enjoyable user experience
5	Clear, visually presented reports that are easy to interpret
6	Modern profiles: 3 dominant factors from the RIASEC model
7	Ipsative scale: participants must prioritise certain factors over others, which limits social desirability bias and prevents high scores across all dimensions
8	Based on the Item Response Theory (IRT) to produce more accurate and reliable results by taking into account the relevance of each question

This manual and user guide will provide useful information about the general conception and technical aspects of **IMMERSION**.

General information - Chapters 1, 2, and 3: Questionnaire development, theoretical background, and practical applications.

→ We recommended these chapters for all candidates and HR practitioners.

Technical information - Chapters 4, 5 and 6: Statistics and psychometric properties.

→ We recommended these chapters for those wanting more insight into the statistical methods used.

1. General Description and Applicability

Every individual tends to seek a work environment that aligns with their identity (Nauta, 2010), which influences their career choices, job satisfaction and professional stability (Bullock et al., 2010). Holland's typology, also known as the RIASEC model, suggests that professional interests are linked to underlying personality types, which help identify the environments best suited to each person (Holland, 1997; Gottfredson, 1999). It distinguishes six core factors, to which each individual relates to varying degrees: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. Each of these dimensions is defined by a specific combination of interests, preferred activities, beliefs, skills, values and personal traits (Nauta, 2010). **IMMERSION** assesses participants' professional interests through these six factors, defined below.



Figure 2. The Structural Framework of Holland's (1977) RIASEC Model of Occupational Themes.

REALISTIC

People with high scores in Realistic interests are 'doers' who like to work on things manually. They are driven by concrete tasks that deliver solid or tangible results or require dexterity and physical movement. Such that, they are likely to prefer activities that involve being outdoors, manual activities and applying practical and technical knowledge. They are likely to be independent and logical individuals who prefer things

to be done in a practical sense. Due to their tendencies and preferences, they are likely to be most comfortable in working environments that allow them to work with or use technical equipment, work on-site, travel, or move around. They tend to dislike being confined to indoor areas such as an office.

Table 2. Description of the REALISTIC sub-factors.

R1	Physical & Outdoor Activities	Preference to engage in physical activities and work outdoors.
R2	Manual & Technical Interests	Preference to make concrete objects and to engage in activities that involve manual work and/or the use of machinery.
Values		Confidence, concrete methods, tangibility and stability.
Learning style		Sensory, experience and doing.

INVESTIGATIVE

People with high scores in Investigative interests are 'thinkers' who are intellectually curious. They are driven by logic and information/facts and are drawn to tasks that require problem-solving and theoretical knowledge. As such, they are likely to prefer activities involving analytical thinking or exploring reasons behind various things to better understand the world around them. Due to their preferences, they are likely to fit better in environments which allow them to continuously learn and develop themselves, as they like to broaden their knowledge and grasp new skills.

Table 3. Description of the INVESTIGATIVE sub-factors.

I1	Intellectual Curiosity & Learning	Preference to engage in activities that require critical thinking and gaining knowledge.
I2	Science & Technology	Preference towards scientific advancements and new technologies, as well as to investigate such topics.
Values		Curiosity, logic, research and learning.
Learning style		Reading, analysing, problem-solving and exploration.

ARTISTIC

People with high scores in Artistic interests are 'creators' and have artistic sensibilities. They value art and like to express themselves through creative means. They are driven by imagination and are drawn to tasks that need authentic designs and individualistic expressions. As such, they are likely to prefer activities that may involve originality and designing things in an appealing way. Due to their tendencies and preferences, they are likely to be most comfortable in working environments which are flexible and encourage freedom to explore different options, allowing them to express their ideas and create new things.

Table 4. Description of the ARTISTIC sub-factors.

A1	Aesthetic sense and expression	Preference towards engaging in activities that require artistic sensitivity and expressing one's emotions in different forms.
A2	Creativity and design	Preference towards conceiving, imagining and innovating things.
	Values	Originality, expression, aesthetics, imagination and individuality.
	Learning style	Design, communication through art, conceptualisation of ideas, and multisensorial environments.

SOCIAL

People with high scores in Social interests enjoy communicating and connecting with other people, and in some cases, they might enjoy providing support and comfort to other people. They are people-driven and are drawn to tasks involving verbal and interpersonal skills and working with people. As such, they are likely to prefer activities that allow them to form personal relationships and work in teams or public-oriented roles. Due to their preferences, they are likely to be most comfortable in working environments which encourage personal contact, collaboration, teamwork and communication.

Table 5. Description of the SOCIAL sub-factors.

S1	Dedication to others	Preference towards helping others, listening, informing, training, giving advice and providing support to the community.
S2	Creating personal relationships	Preference to be in contact with other people, explain, counsel and communicate with clients or co-workers.
	Values	Outgoing, collaboration, socialising, and bonding/connecting.

Learning style	Interaction, sharing and collaborating with others and communicating and listening to others.
-----------------------	---

ENTERPRISING

People with high scores in Enterprising interests like to move things forward, start new projects, and are not shy of assuming risks. They are driven by achieving objectives and drawn to tasks that involve selling their ideas, persuading or leading others. They like to take the initiative and work towards goals that require risk-taking. Due to their interests, they are likely to be most comfortable in working environments which allow them opportunities for being proactive, seizing opportunities, and/or leading others.

Table 6. Description of the ENTERPRISING sub-factors.

E1 Enterprising	Preference towards developing new activities, materialising ideas or projects, taking risks and managing difficult situations.
E2 Leadership	Preference to encourage, motivate and manage or lead a team/other people.
Values	Proactivity, assertiveness, action, risks and competition.
Learning style	Participation, sharing ideas, and experience.

CONVENTIONAL

People with high scores in Conventional interests like tasks that can be carried out in a systematic and methodical way, following an organised process. They are drawn to tasks that involve following instructions in a defined way and might also like to work with numbers. As such, they are likely to prefer activities that may involve analytical thinking and systematic processes. They like to be organised and detail-oriented in their work, following a set structure throughout. Due to their tendencies and preferences, they are likely to be most comfortable in working environments that are structured with well-defined roles and clear expectations, defined processes and regulations.

Table 7. Description of the CONVENTIONAL sub-factors.

C1 Methodical	Preference towards organising, planning tasks, rigorously applying methods or rules and controlling quality or structure.
----------------------	---

C2 Data and numbers	Preference to reason with numbers, analyse statistical data and informatics, as well as understand the laws of mathematics.
Values	Organisation, attention to detail, order, efficiency and preference for numbers.
Learning style	Systematic process, working with numbers/data, standardised rules, guidance and feedback.

1.1. Main Areas of Application

IMMERSION was designed with diversity in mind and is recommended to be taken by all, especially students, graduates and those who want to change their professional career path. It is specifically intended for use by HR professionals and career counsellors; some possible applications can include:

Table 8. Uses of IMMERSION.

Different uses of IMMERSION	
Career Counselling	To help students, trainees, recent university graduates and early career professionals choose a career path.
Career Development	To help professionals explore new occupational avenues and expand their skill base (horizontal career development).
Mentoring/Coaching	To understand the vocational interests of the mentee/coachee and guide them in their career.

It can also be combined with other assessments to gain broader insights into the individual's personality. Some examples of other relevant assessments include:

Table 9. Other applicable assessments.

Assessment name	Advantages
PROFESSIONAL PROFILE 2	In-depth information about interests and personality fit tailored for a specific career.
MOTIVATION+	Information about interests and work motivators that contribute to job satisfaction.

1.2. Questionnaire Structure

The **IMMERSION** questionnaire has been designed to be easily understood and quick to take. The questions are designed using a situational game approach, immersing participants in department-related scenarios.

1.2.1. Format

Unlike traditional questionnaires, questions in **IMMERSION** are presented in a forced choice (FC) format whereby respondents select the answer that best represents their preference from two equally desirable options. This forces respondents to select the option that is most appealing to them.

1.2.2. Questions

The questionnaire consists of 43 questions that measure 6 factors and 12 facets. All questions must be answered for respondents to progress to the next stage.

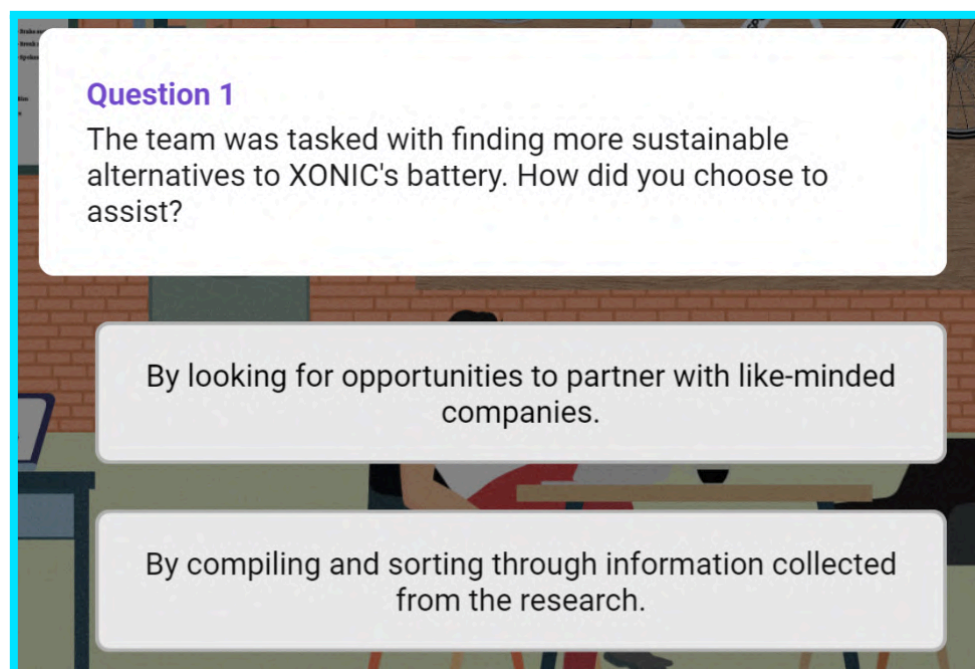


Figure 3. Example of how the questions are presented in the assessment.

1.3. Languages and Special Requirements

1.3.1. Languages available

The questionnaire is currently only available in English, French and Spanish. For more updated information on language availability, please contact the *Central Test Consulting Service* (consulting@centraltest.com).

1.3.2. Accessibility options and special requirements

IMMERSION was designed to be used by individuals from diverse backgrounds and those with a range of special needs. Candidates can adjust display settings such as text size and display speed, full-screen mode, contrast (light/dark mode), and access to instructions at any time during the assessment. Similarly, respondents can access test instructions throughout the assessment at any time. Furthermore, only 1 question is presented per page to help with attentional focus, especially for those who are neurodiverse. Alternatively, for those with visual disabilities, items may be read aloud by a trained administrator. There is no time limit, allowing everyone to complete the assessment at their own pace.

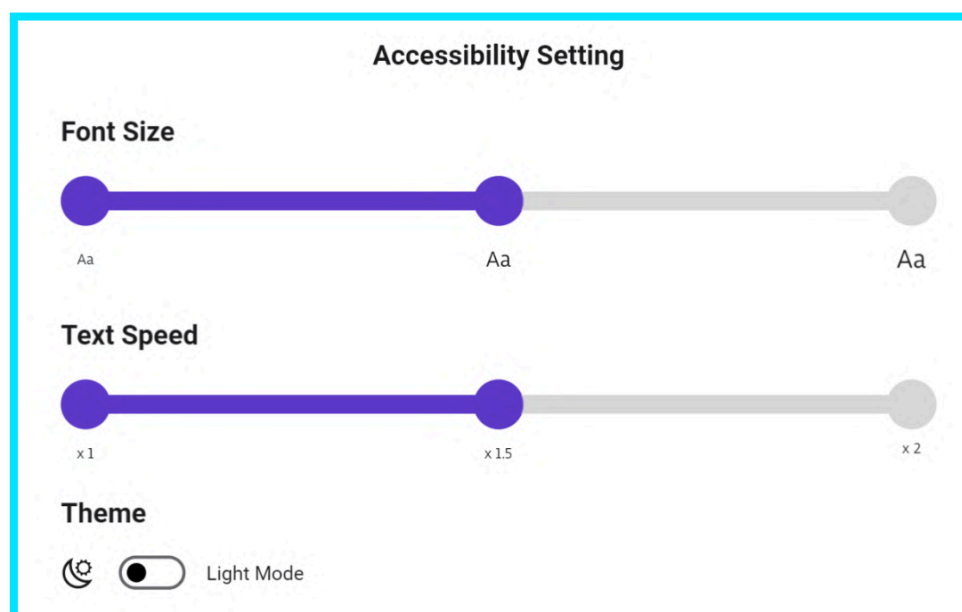


Figure 5. Users can adjust the settings according to their accessibility needs.

1.4. Social Desirability

Social desirability refers to the tendency to present oneself in a way that is viewed favorably by others. When individuals try to show themselves in the best possible light, their results may be biased and not accurately reflect their true functioning (Dodou & de Winter, 2014). To minimise this bias, **IMMERSION** uses an ipsative (forced-choice) questionnaire format. Ipsative measures are more resistant to response manipulation

than Likert scales (Bowen et al., 2002), thereby reducing the possibility for candidates to influence their scores through strategic answering.

1.5. Administration

IMMERSION is administered online and works with all computer operating systems, including Windows, macOS and Linux. The assessment can be administered remotely without supervision, and reports are solely generated online. Standardised instructions are provided to candidates to ensure consistent testing conditions and comparable results. Completion times are recorded to allow administrators to identify individuals who completed the test within an unusually short or long timeframe. Candidates are encouraged to respond spontaneously, focusing on their interests rather than their knowledge or professional experience.

1.6. Technical and Security Issues

The following sections provide recommendations to ensure that the administration of **IMMERSION** is conducted in a secure, objective and fair manner, while safeguarding candidate information during and after the assessment.

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, and the GDPR (General Data Protection Regulation) regulations made by the European Parliament & Council. Assessment results and candidates' personal data may not be shared with third parties without their explicit consent. Candidates also have the right to access, correct, or delete their data at any time by contacting the evaluator or Central Test.

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 encrypted using SSL. Secure browsing is guaranteed by the HTTPS protocol, represented by an icon in the form of a padlock in the Internet

browser.

The application has a « GeoTrust SSL True BusinessID » digital signature, issued and certified by GeoTrust®. This certificate is irrefutable evidence of our serious attitude towards security and protects all assessments from phishing attacks.

Assessment invitations are sent via secure links that are limited in both usage and duration. Assessment duration and candidates' IP addresses are also recorded to ensure traceability and security.

1.6.3. Hardware Architecture and Server Availability

Central Test uses Google Cloud Platform, offering hosting on the same infrastructure that Google uses internally for products such as its search engine. All servers used are located within the European Union and are run on a Unix/Linux type operating system, featuring:

- 6 Compute Engines
- Load Balancing
- High performance of the SSD disk (680,000 IOPS)
- A 2-hour Incident Response Time
- Monitoring and assistance 24/7/365.

Access to the server administration is restricted to an IP address and is password-secured. Data exchanges are secure. *Central Test's* application has been tested on and is compatible with the following browsers:

- Firefox 35 or above
- Google Chrome 30 or above
- Safari 11 or above
- Microsoft Edge 12 or higher

1.6.4. Computer-generated Reports

Reports are generated within minutes after the questionnaire is completed, both within the application and in PDF format. Their content can be customised according to the administrator's needs. Separate reports can also be configured for the candidate, the company, the recruiter, the coach, or the psychologist.

1.7. Supply, Conditions and Availability

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

- Training is mandatory before using or distributing **IMMERSION**.
- **IMMERSION** can be ordered online or by contacting any *Central Test* customer service representative or local partner worldwide (customer@centraltest.com).
- For more information about training, pricing, or products and services related to **IMMERSION**, please contact a local *Central Test* representative or partner.

2. Conceptual Framework and Development

IMMERSION was developed using robust theoretical concepts, practical and relevant work requirements and advanced scientific analyses. All items were piloted and pretested to ensure that assessed work behaviours were contemporary, global and culture-free.

2.1. Theoretical Background

2.1.2. Links between the RIASEC dimensions and personality traits

Significant positive correlations exist between the dimensions of the RIASEC model and the Big Five personality traits (OCEAN: Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism – Costa & McCrae, 1992). Specifically, the Investigative and Artistic factors are associated with Openness, the Social factor with both Extraversion and Agreeableness, the Enterprising factor with Extraversion, and the Conventional factor with Conscientiousness (Barrick et al., 2003; Larson et al., 2002).

2.1.3. Subdivision of the RIASEC factors

The RIASEC dimensions have been subdivided into sub-factors in the Career Insight Questionnaire (CIQ) developed by Dingemanse et al. (2007), allowing for the assessment of more specific categories. For example, Science and Theory as sub-factors of the Investigative dimension. Research has shown that certain correlations between RIASEC and personality traits sometimes apply only to specific sub-factors of a dimension (Dingemanse et al., 2015). In the context of **IMMERSION**, each RIASEC factor has therefore been divided into two sub-factors to better capture the diversity of professional interests.

2.2. Gamification of Psychometric Assessments

Gamification is a learning technique that applies the dynamics of games to other areas, such as professional or educational settings. Initially used to capture students' attention and enhance learning, it later expanded to organisations and businesses, where its benefits have been widely recognised.

2.2.1. Gamified assessments vs. traditional assessments

The main advantage of gamified psychometric tests over traditional tests is the enhancement of user experience (Leutner et al., 2023). The gamified format allows for a more immersive and experiential experience in processes that are often unattractive. Therefore, a better user experience has a positive impact and increases the likelihood of the individual identifying with the situation presented to them and maintaining attention for a longer period, which implies better responses and higher-quality data. They also contribute to a positive image of the organisations that use them, which are perceived as modern, accessible, and attentive to individual well-being.

2.2.2. Management of expectations for gamified assessments

IMMERSION has been psychometrically validated and then gamified. As a new type of tool, gamified assessments generate expectations that might differ for each population group. Unrealistic expectations might lead to negative impacts, e.g., disappointment or low face validity. It is important to identify the attitude of the group it will be used with towards them, and decide if it is the right tool for them, or which information to provide. The negative impact might involve expecting something different, e.g. more interactive and dynamic, or considering that the format does not support enough scientific quality. This information must be transparent and clearly explain the purpose of the assessment, the variables being measured, the format used, and its scientific validity.

2.3. Development of the Tool

IMMERSION was developed from a thorough multi-stage process in a comprehensive and scientific manner. Questionnaire items and factors have been created to ensure that:

- items reflect contemporary ways of working
- effective prevention of faking and social bias
- practitioners have a tool that improves the accuracy of assessment decisions
- individuals are guided when assessing fit with desired roles and careers.
- its structure meets the highest standards of reliability and validity

2.3.1. Factors and Item Construction

IMMERSION was developed from the bank of items used from the original *IMMERSION* questionnaire. The original questionnaire was reviewed and developed in three steps:

Step 1: Feedback was collected from client interviews and training workshop feedback forms to ensure that the questionnaire and the measured factors meet their needs.

Step 2: Review of items by a group of psychometric experts, with particular attention to vocabulary clarity and cultural neutrality.

Step 3: The revised questionnaire was then pretested three times with a representative sample of people. Results were analysed, and items were removed and/or changed based on their psychometric qualities, to achieve the best reliability and validity.

2.3.2. Combined Profiles

Most RIASEC tools establish an individual's professional interest profile by combining the most dominant factors among the six dimensions of the model. As such, **IMMERSION** reports present a profile based on the combination of the top three interests, ranked from strongest to weakest. For example, for someone with an EIC combined profile:

- Enterprising = Most dominant interest
- Investigative = 2nd main interest
- Conventional = 3rd main interest

A match with Holland's profiles defined by the combination of the two highest interests (regardless of their relative strength) is also provided.

2.3.3. Ipsative Response Format

Response options for the items in **IMMERSION** are written in binary forced-choice format (FC). This response format bypasses several limitations of the Likert scale (Brown & Maydeu-Olivares, 2011a), enhances score differentiation, and provides a more accurate representation of individual preferences (Cruise, 2011/2012). As a result, it reduces response biases and social desirability while improving the reproducibility of results.

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

The combination of Item Response Theory (IRT) and Thurstonian comparative models (Thurstone, 1927; 1931) helps overcome certain limitations of ipsative data and corrects distortions that can affect the validity of results. This approach was used to revise the items and factors in **IMMERSION**.

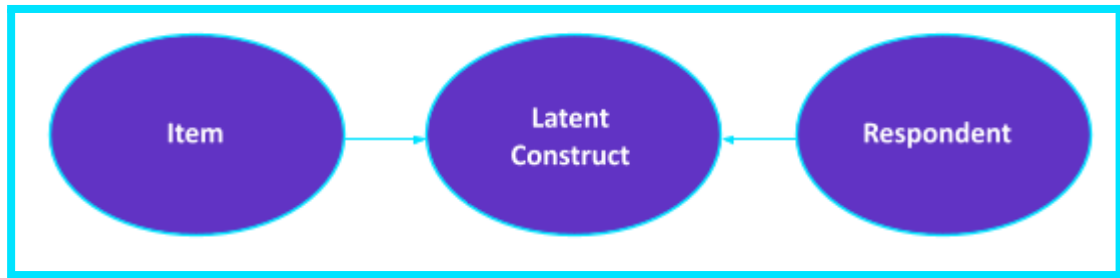


Figure 5. Model of Thurstonian IRT Analyses Used in the Development of **IMMERSION**.

2.3.5. Item Difficulty and Discrimination

Item difficulty and discrimination were assessed in the validation of items:

Item difficulty → the level of interest at which a participant in the sample has a 50% chance of selecting a given response. Thresholds were established to identify items where one response was systematically favoured.

Item discrimination → an item's ability to differentiate individuals based on their level of interest. Items with very low discrimination were excluded or revised.

2.3.6. Pretests and Pilot Study

Pretests were conducted with a range of male and female participants from different countries to ensure that items were accurate, relevant, clear, concise, culture-free, bias-free and reflected targeted work behaviours. The pretest data were collected from the Testing Center, the Test&Go website (*Central Test*), and CINT (a sampling firm). Participants who took the assessment from the Testing Center and Test&Go received a full report as compensation for their participation, whilst participants from CINT received monetary compensation. A final pretest conducted with native English speakers from various continents confirmed the questionnaire's international validity for assessing professional interests.

3. Scoring

3.1. Raw Scores

A raw score is the total number of responses endorsed by a candidate for each factor measured. Since each IMMERSION item is scored in a binary way (0 = not chosen, 1 = chosen), the raw score is the sum of all chosen responses within a factor. To ensure fairness and standardisation, raw scores are converted into standard scores on a 0–10 scale.

3.2. Conversion of Raw Scores to Standard Scores

For each **IMMERSION** factor, the maximum possible raw score depends on the number of items included under each factor, while the minimum possible raw score is always 0.

The conversion of raw scores is carried out in a two-step process:

Step 1: Ipsative scores are calculated for each scale by summing all the forced-choice responses for each factor. In each response, the chosen option is marked with a 1 and the rejected option with a 0.

Step 2: The scores are then transformed to a 0-to-10 scale using the following formula:

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

Where:

- max is the maximum possible score for each factor
- min is the minimum possible score for each factor

3.3. Scoring in the Report

IMMERSION reports reflect the scoring process, with each individual receiving a score from 0 to 10. The report provides a detailed description of each of the six RIASEC factors assessed in the assessment. The level corresponding to the candidate's score is selected from six possible levels, detailing their preferences and the activities or occupations they are most inclined toward.

Higher scores reflect stronger intensity, while lower scores reflect weaker intensity. For example, a candidate with a low score in a particular factor might be described as “does not seem to show much interest,” whereas a high score might indicate “a strong interest” in the relevant factor. Scores should be interpreted in relative terms, with the

ranking of profiles being more important than the absolute score.

The combined profiles in the report also reflect the scoring process. These profiles represent the combination of the three factors with the highest scores. They are calculated by considering that the first two interest areas have the highest and equal importance, while the third interest area carries slightly less weight.

4. Psychometric Properties

4.1. Sample Descriptives

The sample comprised $N = 306$ professionals from English-speaking countries ($n = 159$ females, $n = 144$ males, $n = 3$ undisclosed), ranging in age from 18 to 77 years ($M=39.6$, $SD=16.70$). All the participants represented a range of occupational groups. Data was collected from the Testing Centre and a sampling firm named CINT, and each respondent received compensation for their participation. To ensure the quality of data obtained, participants whose answers were deemed unreliable or who completed the questionnaire in less than 8 minutes were excluded from analyses.

Tables 10-12 provide an overview of the sample descriptives.

Table 10. *Country of Origin*

Candidate Country	Percentage (%)	N
United Kingdom	34	104
Canda	24	73
USA	13	40
Other	27.7	85
N/A	1.3	4

Table 11. *Gender*

Gender	Percentage (%)	N
Female	51.96	159
Male	47.06	144
N/A	0.98	3

Table 12. *Age Groups*

Age Group	Percentage (%)	N
18 to 25	18.30	56

26 to 33	20.92	64
34 to 41	18.95	58
42 to 49	14.38	44
50 and above	24.84	76
N/A	2.61	8

4.2. Descriptive Norms

The normative population refers to the larger group that the assessment is intended to represent. It is defined by collecting data from a sufficiently large number of candidates who complete the assessment. From these data, descriptive statistics are estimated that are generalised to reflect what would be expected in the broader population. In this way, the normative population serves as the statistical reference point for interpreting individual scores.

For each RIASEC factor and facet, the results are summarised through descriptive norms. These include the mean (average score), minimum (lowest observed score), maximum (highest observed score), standard deviation (degree of variability), range (spread of scores from minimum to maximum), and the standard error (precision of the mean estimate). Together, these indices provide an empirical context for interpretation. For example, if the mean score for a factor is 4.5, it shows that people in the normative population typically score around 4.5 on that factor. In this way, descriptive norms establish the statistical framework in which individual results can be located and evaluated relative to the broader population distribution.

Tables 13-18 detail the mean (*M*), standard deviation (*SD*), minimum (Min), maximum (Max), range, and standard error (*SE*) scores on factors and subfactors of **IMMERSION**, for the whole sample, for the different genders and age groups.

Table 13. *Factor Score Norms for the Whole Sample*

Factor	Mean	SD	Min	Max	Range	SE
Whole sample (N = 306)						
R	4.48	1.43	1.3	8.65	7.35	0.08
I	4.09	1.21	0.55	7.25	6.7	0.07
A	3.77	1.31	0	7.25	7.25	0.07

S	4.57	1.3	0.55	8.6	8.05	0.07
E	4.02	1.46	0.55	8.7	8.15	0.08
C	3.82	1.39	0	7.85	7.85	0.08

Table 14. *Sub-factor Score Norms for the Whole Sample*

Sub-factor	Mean	SD	Min	Max	Range	SE
Whole sample (N = 306)						
R1	5.02	1.83	0	9.3	9.3	0.1
R2	3.95	2	0	9.3	9.3	0.11
I1	4.31	1.74	0	8.9	8.9	0.1
I2	3.87	1.53	0	7.8	7.8	0.09
A1	3.62	1.7	0	8.6	8.6	0.1
A2	3.92	1.62	0	7.8	7.8	0.09
S1	4.74	1.62	0	10	10	0.09
S2	4.4	1.65	0	8.6	8.6	0.09
E1	4.42	1.91	0	8.9	8.9	0.11
E2	3.61	1.83	0	8.6	8.6	0.1
C1	3.97	1.59	0	8.6	8.6	0.09
C2	3.68	1.93	0	8.6	8.6	0.11

Table 15. *Factor Score Norms by Gender*

Factor	Mean	SD	Min	Max	Range	SE
Female (n = 159)						
R	4.53	1.37	1.3	8	6.7	0.10
I	4.24	1.21	0.55	7.25	6.7	0.09
A	3.81	1.38	0	7.15	7.15	0.10
S	4.59	1.36	0.55	8.6	8.05	0.10
E	3.96	1.49	0.55	8.7	8.15	0.11

C	3.77	1.39	0	7.25	7.25	0.11
Male (n = 144)						
R	4.46	1.48	1.35	8.65	7.3	0.12
I	3.92	1.19	0.55	7.25	6.7	0.09
A	3.73	1.25	0.55	7.25	6.7	0.10
S	4.55	1.23	1.1	7.25	6.15	0.10
E	4.07	1.42	0.55	7.25	6.7	0.11
C	3.89	1.38	0	7.85	7.85	0.11

Table 16. Sub-factor Score Norms by Gender

Sub-factor	Mean	SD	Min	Max	Range	SE
Female (n = 159)						
R1	5.14	1.88	0	9.3	9.3	0.14
R2	3.92	1.98	0	8.6	8.6	0.15
I1	4.56	1.81	0	8.9	8.9	0.14
I2	3.92	1.48	0	7.8	7.8	0.11
A1	3.71	1.78	0	8.6	8.6	0.14
A2	3.91	1.63	0	7.1	7.1	0.12
S1	4.79	1.63	0	10	10	0.13
S2	4.40	1.76	0	8.6	8.6	0.13
E1	4.32	1.97	0	8.8	8.8	0.15
E2	3.59	1.84	0	8.6	8.6	0.14
C1	3.95	1.66	0	7.8	7.8	0.13
C2	3.59	1.92	0	8.6	8.6	0.15
Male (n = 144)						
R1	4.92	1.77	0	9.3	9.3	0.14
R2	4.00	2.00	0	9.3	9.3	0.16
I1	4.03	1.61	0	7.8	7.8	0.13
I2	3.81	1.59	0	7.8	7.8	0.13

A1	3.50	1.61	0	7.8	7.8	0.13
A2	3.96	1.62	0	7.8	7.8	0.13
S1	4.70	1.61	1.1	7.8	6.7	0.13
S2	4.40	1.54	1.1	7.8	6.7	0.12
E1	4.53	1.85	1.1	8.9	7.8	0.15
E2	3.61	1.81	0	7.8	7.8	0.15
C1	4.02	1.52	0	8.6	8.6	0.12
C2	3.76	1.92	0	7.8	7.8	0.16

Table 17. *Factor Score Norms by Age Groups*

Factor	Mean	SD	Min	Max	Range	SE
18 to 25 (n = 56)						
R	4.48	1.52	2.00	8.00	6.00	0.20
I	4.00	1.11	1.10	7.25	6.15	0.15
A	4.08	1.38	0.70	7.15	6.45	0.18
S	4.43	1.58	0.55	7.85	7.30	0.21
E	4.46	1.49	1.65	8.70	7.05	0.20
C	3.90	1.45	0.00	7.85	7.85	0.19
26 to 33 (n = 64)						
R	4.67	1.24	1.35	8.65	7.30	0.16
I	4.19	1.33	0.55	6.70	6.15	0.17
A	3.85	1.42	1.10	6.15	5.05	0.18
S	4.29	1.19	1.65	6.70	5.05	0.15
E	3.92	1.26	1.10	7.25	6.15	0.16
C	3.69	1.31	1.10	6.70	5.60	0.16
34 to 41 (n = 58)						
R	4.56	1.38	1.30	6.70	5.40	0.18
I	3.78	1.12	1.10	6.70	5.60	0.15
A	3.72	1.29	0.00	6.15	6.15	0.17
S	4.57	1.21	1.10	7.25	6.15	0.16

E	4.12	1.46	1.10	7.25	6.15	0.19
C	3.84	1.29	1.65	6.70	5.05	0.17
42 to 49 (n = 44)						
R	4.50	1.48	2.00	7.35	5.35	0.22
I	4.18	1.27	1.10	7.25	6.15	0.19
A	3.91	1.26	1.65	7.25	5.60	0.19
S	4.50	1.34	1.10	7.25	6.15	0.20
E	3.64	1.53	0.55	6.15	5.60	0.23
C	3.85	1.69	0.00	7.25	7.25	0.25
50 and above (n = 76)						
R	4.33	1.49	1.30	8.65	7.35	0.17
I	4.25	1.20	1.10	7.25	6.15	0.14
A	3.44	1.19	0.55	6.70	6.15	0.14
S	4.93	1.14	1.65	8.60	6.95	0.13
E	3.88	1.49	0.55	7.25	6.70	0.17
C	3.86	1.34	0.70	6.70	6.00	0.15

Table 18. *Sub-factor Score Norms by Age Groups*

Sub-factor	Mean	SD	Min	Max	Range	SE
18 to 25 (n = 56)						
R1	4.86	1.97	1.30	9.30	8.00	0.26
R2	4.10	1.98	0.00	8.60	8.60	0.26
I1	3.99	1.91	0.00	8.80	8.80	0.25
I2	4.02	1.60	0.00	7.10	7.10	0.21
A1	3.87	1.89	1.10	8.60	7.50	0.25
A2	4.29	1.69	0.00	7.80	7.80	0.23
S1	4.47	1.90	0.00	10.00	10.00	0.25
S2	4.38	1.88	0.00	7.10	7.10	0.25
E1	4.92	1.89	1.10	8.80	7.70	0.25
E2	4.00	2.00	0.00	8.60	8.60	0.27

C1	3.89	1.69	0.00	8.60	8.60	0.23
C2	3.90	1.98	0.00	8.60	8.60	0.26
26 to 33 (n = 64)						
R1	4.93	1.71	0.00	8.00	8.00	0.21
R2	4.41	1.94	0.00	9.30	9.30	0.24
I1	4.23	1.80	0.00	7.80	7.80	0.23
I2	4.16	1.49	1.10	6.70	5.60	0.19
A1	3.49	1.82	0.00	7.80	7.80	0.23
A2	4.21	1.66	0.00	6.70	6.70	0.21
S1	4.67	1.52	1.10	7.80	6.70	0.19
S2	3.91	1.54	0.00	6.70	6.70	0.19
E1	4.49	1.72	1.10	7.80	6.70	0.22
E2	3.34	1.79	0.00	7.80	7.80	0.22
C1	3.67	1.63	0.00	6.70	6.70	0.20
C2	3.72	1.84	0.00	7.80	7.80	0.23
34 to 41 (n = 58)						
R1	5.06	1.54	1.30	8.00	6.70	0.20
R2	4.07	2.12	0.00	8.00	8.00	0.28
I1	3.87	1.43	0.00	7.80	7.80	0.19
I2	3.68	1.53	1.10	7.80	6.70	0.20
A1	3.43	1.71	0.00	6.70	6.70	0.22
A2	4.01	1.64	0.00	6.70	6.70	0.22
S1	4.73	1.67	1.10	7.80	6.70	0.22
S2	4.42	1.59	1.10	7.80	6.70	0.21
E1	4.50	1.92	1.10	8.90	7.80	0.25
E2	3.74	1.72	0.00	6.70	6.70	0.23
C1	4.20	1.52	1.10	7.80	6.70	0.20
C2	3.48	1.88	0.00	6.70	6.70	0.25
42 to 49 (n = 44)						
R1	5.25	1.92	0.00	8.00	8.00	0.29

R2	3.75	1.89	1.30	8.00	6.70	0.29
I1	4.69	1.86	1.10	8.90	7.80	0.28
I2	3.67	1.56	1.10	6.70	5.60	0.24
A1	3.70	1.70	1.10	7.80	6.70	0.26
A2	4.13	1.48	1.10	7.80	6.70	0.22
S1	4.62	1.61	1.10	7.80	6.70	0.24
S2	4.38	1.70	1.10	7.80	6.70	0.26
E1	3.83	1.91	0.00	6.70	6.70	0.29
E2	3.45	1.78	0.00	6.70	6.70	0.27
C1	4.08	1.74	0.00	7.80	7.80	0.26
C2	3.63	2.14	0.00	6.70	6.70	0.32
50 and above (n = 76)						
R1	5.18	2.01	1.30	9.30	8.00	0.23
R2	3.48	1.89	0.00	8.00	8.00	0.22
I1	4.74	1.62	1.10	8.90	7.80	0.19
I2	3.76	1.52	0.00	7.80	7.80	0.17
A1	3.58	1.49	0.00	6.70	6.70	0.17
A2	3.29	1.46	0.00	6.70	6.70	0.17
S1	5.10	1.45	1.10	8.60	7.50	0.17
S2	4.77	1.50	1.10	8.60	7.50	0.17
E1	4.25	1.99	0.00	8.90	8.90	0.23
E2	3.52	1.82	0.00	7.80	7.80	0.21
C1	4.08	1.48	1.10	7.80	6.70	0.17
C2	3.64	1.90	0.00	7.80	7.80	0.22

4.3. Reliability

Reliability refers to the extent to which an assessment is free from error and the degree to which items on a scale measure the same construct. 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, interest, attitude or aptitude. 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 the 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 by errors in human judgment during the recruitment and selection process.

4.3.1. 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's alpha, while a popular way of determining a measure's 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.

Since **IMMERSION** has an ipsative, binary-choice structure, 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 more desirable. A higher CR value indicates that a greater proportion of the variance in item responses is attributable to the underlying construct, rather than to measurement error. As a rule of thumb, a reliability of 0.70 or higher is recommended (Nunnally, 1978), with values greater than 0.90 (excellent); 0.70 to 0.90 (good); 0.60 to 0.70 (acceptable); 0.50 to 0.60 (poor); and values less than 0.50 (unacceptable)

The formula for Composite Reliability (CR) is:

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

The composite reliability for **IMMERSION** was computed on the whole sample ($N = 306$). Table 19 describes the composite reliability of **IMMERSION**. All factors showed composite reliability that fell within the acceptable to excellent range (0.68 – 0.91). This demonstrates a high degree of internal consistency, meaning that the items within each scale reliably measure the same underlying vocational trait. In addition, **IMMERSION** had an overall **composite reliability of 0.80**, indicating that the scale has very good internal consistency and is highly reliable.

Table 19. *Composite Reliability for Factors*

FACTOR	CR
R	0.76
I	0.91
A	0.87
S	0.75
E	0.68
C	0.81

4.4. 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. IMMERSION was validated following a thorough multi-stage process that ensured items and factors were accurate, representative, consistent, clear, fair and culturally free.

4.4.1. 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) developed with insufficient research. Therefore, in order to develop a scale that measures occupational and vocational preferences, the construct must reflect known theories that define interests, identify the components that it is composed of, and pinpoint positive and negative attributes associated with each dimension.

For interest-based assessments, construct validity is important as it: (a) adds legitimacy and acceptability to the standardised measurement of employee interests, (b) facilitates understanding of organisational implications of findings, and (c) enables organisations and individuals to tailor interventions to meet desired targets. Construct validity addresses the question, “Does **IMMERSION** define and classify occupational and vocational preferences in a way that represents theoretical and practical understanding?” For **IMMERSION**, construct validity was established using two approaches: **inter-factor correlations** and **confirmatory factor analysis**, as detailed in the sections below.

4.4.1.a. Inter-factor Correlations

The factors in **IMMERSION** measure occupational and vocational preferences; however, they all measure different facets of interests. In other words, if someone scores highly on a factor, it does not necessarily mean that they will score highly on all factors. Table 20 shows the inter-factor correlations. The correlation coefficient is the numerical way of quantifying the relationship between two variables and ranges from -1 to 1. The closer the value is to 1, the greater the relationship. The correlation coefficient does not imply causality; it simply points out the relationship between two variables, but it is important to consider that there may be explanatory third variables.

Table 20. *Inter-factor Correlations*

	R	I	A	S	E	C
R	1.00					
I	-0.15**	1.00				

A	0.06	-0.23***	1.00			
S	-0.16**	-0.26***	-0.14*	1.00		
E	-0.32***	-0.26***	-0.18**	-0.06	1.00	
C	-0.25***	0.04	-0.32***	-0.27***	-0.17**	1.00

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

The correlation analysis of the RIASEC dimensions provides evidence that is broadly consistent with the circumplex structure proposed by Holland's model. As expected for opposite types, we observe clear negative correlations such as Realistic–Social ($r = -0.16$), Investigative–Enterprising ($r = -0.26$), and Artistic–Conventional ($r = -0.38$). These associations highlight the theoretical contrasts between interest areas positioned across from each other in the hexagon. In addition, some dimensions located two steps apart show weaker or near-zero associations, for example, Realistic–Artistic ($r = 0.06$) and Investigative–Conventional ($r = 0.04$), which is consistent with the expectation that more distant types overlap less strongly.

We also observed some deviations from the expected pattern. However, it should be noted that the ipsative scoring format of the immersion naturally constrains the correlations between scales. In such formats, where choices on one dimension reduce the available score on others, negative intercorrelations are mathematically induced, even between adjacent dimensions that would typically be expected to correlate positively in a normative measure. The use of Thurstonian IRT (TIRT), however, allows us to recover normative estimates from these data and provides a more accurate representation of the RIASEC structure.

These deviations are not unique to our study: other researchers have also observed that RIASEC intercorrelations often diverge from the ideal circumplex, especially when forced-choice or ipsative scoring is used (e.g., Schulte et al., 2020; Rasheed et al., 2022). Such findings reinforce the importance of advanced modelling approaches like TIRT to mitigate structural distortions and to better approximate the theoretical RIASEC model (Brown & Maydeu-Olivares, 2013).

4.4.1.b. Confirmatory Factor Analysis

A confirmatory factor analysis (CFA) with continuous factor indicators was conducted to check the convergence of the results with the assessment model. CFA is a form of factor analysis used to measure whether a construct is consistent with its intended meaning.

IMMERSION uses a strict confirmatory approach, where 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.

CFA is used to determine how well the model fits the data, and therefore, evaluating the goodness of fit indices is crucial. For the validation of **IMMERSION**, 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). Like the χ^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 the overall improvement of a proposed model compared with an independent 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 an RMSEA of .05 or less (Hu & Bentler, 1999).

Thurstonian IRT (TIRT) based CFA

IMMERSION is based on a binary ipsative forced-choice structure, which violates many of the assumptions of traditional CFA based on Likert-style responses. To address this, the Thurstonian Item Response Theory (TIRT) model is used. Item response theory (IRT) refers to a set of mathematical models that describe, in probabilistic terms, the relationship between a person's response to a survey question and his or her level of the 'latent variable' being measured by the scale (Hays & Fayers, 2005, p.55). In ipsative tests, a respondent must make a comparative judgement between two or more items to give an answer. Conversely, in normative tests, judgment is measured in absolute terms for each evaluated attribute and does not compare preferences between different attributes. Classical test scoring does not reflect this psychological process of comparison, but rather ignores the fact that respondents indicate a preference, not an absolute judgement. Thurstone's law of comparative judgments (Thurstone, 1927; 1931) generates results based on the comparative judgments of individuals.

TIRT is specifically designed for forced-choice assessments, where respondents choose between options rather than rate them independently. It accounts for the comparative nature of responses, dependencies between item pairs, and the latent structure underlying choice data, making it more appropriate and psychometrically robust for this

type of assessment. TIRT integrates with the CFA framework by modelling the latent traits that govern response behaviour across pairwise items. Instead of treating each item as independent, TIRT models each pair of items (blocks) as a function of underlying trait differences. This ensures more accurate parameter estimation (e.g., factor loadings, trait estimates) as well as allows us to assess model fit using conventional SEM indices such as RMSEA, CFI, and TLI.

In summary, the TIRT-based CFA framework provides a rigorous and reliable method for evaluating IMMERSION's measurement model, making it especially suited for binary ipsative, RIASEC-based career assessments like **IMMERSION**.

In order to analyse the data using Thurstonian IRT modelling, raw data were converted into binary outcomes following the procedure described by Brown and Maydeu-Olivares (2011) for Item response modelling of paired comparisons and ranking data. The observed responses to the paired comparisons were coded using the formula below, where l indicates the pair $\{a, b\}$;

$$y_l = \begin{cases} 1 & \text{if item } a \text{ is preferred over item } b \\ 0 & \text{if item } b \text{ is preferred over item } a \end{cases}$$

The TIRT model was applied to the binary outcomes of paired comparisons, with models specified at both the 6-factor (RIASEC) and 12-subfactor levels. In total, 43 binary outcomes were generated from 43 blocks of two items. The models were estimated in R using the Thurstonian IRT package (Thissen & Brown, 2019) in conjunction with *lavaan* (Rosseel, 2012). Estimation was carried out with the unweighted least squares mean and variance adjusted (ULSMV) algorithm, which is appropriate for binary forced-choice data and provides robust model fit indices.

Tables 21 and 22 show the goodness of fit indices and model results.

Table 21. *Confirmatory Factor Analysis using Thurstonian IRT for the 6 Factor Model*

N	χ^2	df	p	RMSEA	$p \leq .05$	CFI	TLI
306	958.04	802	0	0.02	1	0.71	0.67

Table 22. *Confirmatory Factor Analysis using Thurstonian IRT for the 12 Factor Model*

N	χ^2	df	p	RMSEA	$p \leq .05$	CFI	TLI
---	----------	----	---	-------	--------------	-----	-----

306	846.23	751	0	0.02	1	0.82	0.78
-----	--------	-----	---	------	---	------	------

As reported in Tables 21 and 22, the RMSEA values were 0.02 and 0.02 for 6 and 12-factor models, respectively. These values fall well below the commonly accepted cutoff of 0.05, which indicates a close fit of the model to the data (Steiger & Lind, 1980). An RMSEA of 0.02 is considered excellent, as values below 0.05 are generally interpreted as reflecting a close fit, and values below 0.08 as an acceptable fit. Thus, the RMSEA values obtained for both models indicate that the gap between the model's expected structure and the observed data is very small. In terms of model evaluation, this indicates that the specified TIRT models provide an excellent approximation to the data and can be considered well-fitting representations of the underlying structure. The 12-factor model demonstrated comparatively stronger incremental fit, with higher CFI (0.82) and TLI (0.78) values relative to the 6-factor solution, which gives more credence to the idea that each factor in the RIASEC model can be split into two facets. Therefore, the latent dimensions captured by the **IMMERSION** assessment align closely with the expected RIASEC structure, supporting the validity of both the broad six-factor RIASEC model and its finer twelve-subfactor extension.

5. IMMERSION and Equal Opportunities

Organisations are increasingly pressured 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. **IMMERSION** is a fair and accurate measure of occupational and vocational preferences, 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 in 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)

5.1. 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 of a trait have an equal chance of being selected for a role that requires the desired level of this trait. ME helps to determine the extent to which respondents are interpreting items similarly and identifies differences across groups that reflect actual differences in the trait being measured (Drasgow, 1984). One approach to assessing measurement equivalence is differential item functioning analysis.

5.1.1. 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). Here, the logistic regression approach is used, whereby separate analyses are run for each item, and maximum likelihood estimation is used to calculate the probability of endorsing an item.

5.1.1.a. DIF Analysis: Gender

DIF analysis was performed across gender in order to ensure that there was not a bias towards either males or females, and therefore that **IMMERSION** is a fair test for both genders. For this analysis, the whole sample was divided according to gender: 159 females and 144 males.

Table 23. *Coefficients for Significant DIF (Gender) and Effect Size for DIF Items*

Question	Coefficient	R ²	Effect size
R2	8.422	0.04	Negligible
I1	6.905	0.03	Negligible
I1	7.107	0.03	Negligible
S1	8.422	0.04	Negligible
S2	7.107	0.03	Negligible
E1	6.351	0.03	Negligible

C1	6.179	0.03	Negligible
C2	6.905	0.03	Negligible
C1	6.351	0.03	Negligible

The analysis identified 9 out of the 86 items with a DIF effect for gender. Table 23 shows the DIF coefficient as well as the R^2 effect size. According to Zumbo and Thomas' (1997) classification of effect size, any R^2 under 0.13 is considered negligible. Out of the 10 items flagged for DIF 0 showed an effect size of 0.13 or over. Therefore, it can be considered that **IMMERSION** does not contain any item that functions differently across genders

5.1.1.b. DIF Analysis: Age

DIF analysis was performed based on age in order to ensure that **IMMERSION** was a fair test for people of all ages. For this analysis, two age groups were defined: 222 aged between 18 and 49 ($M=32.53$; $SD=8.34$), and 76 aged 50 or over ($M=63.3$; $SD=9.11$).

Table 24. *Coefficients for Significant DIF (Age) and Effect Size for DIF Items*

Question	Coefficient	R^2	Effect size
R1	5.86	0.03	Negligible
S1	7.38	0.03	Negligible
E2	5.86	0.03	Negligible
E1	7.38	0.03	Negligible

The analysis flagged 4 of the 86 items for DIF based on our age groups. However, as demonstrated in Table 24, the effect size for each of these items is negligible. Therefore, it can be concluded that these items are not discriminatory against people of an older age.

5.1.2. T-tests and ANOVAs

The DIF analysis demonstrates that **IMMERSION** has high consequential validity, and although several items were flagged with DIF, there was a negligible impact on the

overall assessment, meaning that the assessment does not bias towards gender or age. Therefore, any significant differences found between genders or age groups can be attributed to real differences between the groups and not biases of the assessment. To further explore the differences between the groups, t-tests and ANOVAS were performed between genders and age groups (see Tables 25 to 30).

5.1.2.a. Age Groups by Factors

Participants were separated into five groups according to their age: 18 to 25 ($n=56$), 26 to 33 ($n=64$), 34 to 41 ($n=58$), 42 to 49 ($n=44$), 50 and above ($n=76$).

Table 25. ANOVA for Factors by Age Group

	18 to 25	26 to 33	34 to 41	42 to 49	50 and above	F	df	p-value
R	4.48	4.67	4.56	4.50	4.33	0.55	4	0.7
I	4.00	4.19	3.78	4.18	4.25	1.58	4	0.18
A	4.08	3.85	3.72	3.91	3.44	2.24	4	0.07
S	4.43	4.29	4.57	4.50	4.93	2.49	4	0.04*
E	4.46	3.92	4.12	3.64	3.88	2.38	4	0.05*
C	3.90	3.69	3.84	3.85	3.86	0.20	4	0.94

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

Table 26. Post-hoc T-tests for Factors by Age Groups.

Age Group	18 to 25	26 to 33	34 to 41	42 to 49	50 and above
S					
18 to 25	-				
26 to 33	0.58	-			
34 to 41	0.57	0.19	-		
42 to 49	0.80	0.38	0.77	-	
50 and above	0.03*	0.00**	0.08	0.06	-
E					
18 to 25	-				
26 to 33	0.03	-			

34 to 41	0.22	0.40	-		
42 to 49	0.00**	0.30	0.10	-	
50 to 56	0.02*	0.88	0.35	0.39	-

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

The ANOVA for the RIASEC factors by age groups found significant differences for S and E. Post hoc analysis was conducted using pairwise t-tests in order to see where these differences occurred.

For Factor S, the age group 50 and above ($M=4.93$) was significantly higher than the age group 18-25 ($M=4.43$) and 26 to 33 ($M = 4.29$). The results indicate that the old group scored higher on the Social (S) and Enterprising (E) factors compared to the younger groups. This may be because, with age, people tend to place greater importance on meaningful social relationships, mentoring, and community involvement, which raises their interest in social activities. At the same time, older adults are more likely to take on leadership or advisory roles at work, supported by their greater experience, confidence, and networks.

5.1.2.b. Age Groups by Sub-factors

Table 27. ANOVA for Age Groups by Sub-factors

Sub-factor	18 to 25	26 to 33	34 to 41	42 to 49	50 to 56	F	df	p-value
R1	4.86	4.93	5.06	5.25	5.18	0.42	4	0.79
R2	4.1	4.41	4.07	3.75	3.48	2.24	4	0.07
I1	3.99	4.23	3.87	4.69	4.74	3.24	4	0.01**
I2	4.02	4.16	3.68	3.67	3.76	1.18	4	0.32
A1	3.87	3.49	3.43	3.7	3.58	0.58	4	0.68
A2	4.29	4.21	4.01	4.13	3.29	4.49	4	0***
S1	4.47	4.67	4.73	4.62	5.1	1.4	4	0.23
S2	4.38	3.91	4.42	4.38	4.77	2.43	4	0.05*
E1	4.92	4.49	4.5	3.83	4.25	2.25	4	0.06
E2	4	3.34	3.74	3.45	3.52	1.2	4	0.31
C1	3.89	3.67	4.2	4.08	4.08	1.05	4	0.38

C2	3.9	3.72	3.48	3.63	3.64	0.36	4	0.84
-----------	-----	------	------	------	------	------	---	------

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

Table 28. *Post-hoc T-Tests for Sub-factors by Age*

Age Group	18 to 25	26 to 33	34 to 41	42 to 49	50 to 56
I1					
18 to 25	-				
26 to 33	0.4759	-			
34 to 41	0.7073	0.226	-		
42 to 49	0.0663	0.1976	0.0129**	-	
50 and above	0.0152*	0.0776	0.0014**	0.877	-
A2					
18 to 25	-				
26 to 33	0.7896	-			
34 to 41	0.3677	0.5037	-		
42 to 49	0.6048	0.783	0.7167	-	
50 and above	0.0004***	0.0007***	0.0086*	0.0034**	-
S2					
18 to 25	-				
26 to 33	0.1315	-			
34 to 41	0.91	0.0735	-		
42 to 49	1	0.1313	0.9195	-	
50 and above	0.1898	0.001**	0.1926	0.2003	-

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

For an in-depth look at the differences between age groups for the RIASEC factors, ANOVA and t-tests were conducted on the sub-factors. The results are shown in Tables 27 and 28. The differences found for Social were shown to exist in only S2 and not S1. The post hoc t-test result for S2 shows that the 26 to 33 group ($M=3.91$) were both significantly lower than the 50 and above group ($M=4.77$). This indicates that older adults may especially value the aspects of Social captured by S2, such as building and

maintaining personal relationships, and communicating with others in their work and social life. In contrast, adults in the 26 to 33 age range may be more focused on career development or personal achievement, which could explain their lower scores on this facet of Social.

The differences found for Enterprising neither exist in E1 nor E2, suggesting that the overall effect for Enterprising arises from small but consistent differences across both sub-factors that, when combined, yield a significant effect at the higher-order factor level. This indicates that the difference with age may not appear strongly in one single aspect of Enterprising, but rather emerge when considering the broader construct. This also highlights the importance of examining both factor-level and sub-factor-level scores, as effects may emerge more clearly at one level of aggregation than another.

The post hoc t-tests for A2 indicate that the 50 and above group ($M=3.29$) was significantly lower than all other age groups: 18 to 25 ($M=4.29$), 26 to 33 ($M=4.21$), 34 to 41 ($M=4.01$), 42 to 49 ($M=4.12$). This indicates that older adults show less interest in the aspects of A2, which involve creativity, artistic expression, and imaginative engagement, compared to younger adults. This difference could be due to the fact that A2 is more centred around creativity and thinking outside the box, traits more often seen in younger generations who have a fresh outlook on the world and greater exposure to creative fields. Perhaps the reason why this difference is not also seen in A1 is that A1 relates to aesthetic appreciation, which is far less likely to change with age.

The post hoc t-tests result for I1 shows that the 50 and above group ($M=4.74$) scored significantly higher than the 18 to 25 group ($M=3.99$) and the 34 to 41 group ($M=3.87$), while the 42-49 group ($M=4.69$) was significantly higher than the 34 to 41 group ($M=3.87$). This indicates that older adults, particularly those aged 50 and above, may place greater importance on seeking knowledge and reflecting on ideas. Rather than focusing only on career-building, older adults may be more motivated to engage with knowledge for personal growth and enrichment, drawing on their broader life experiences, compared to younger adults who may be more focused on early career development or immediate practical goals.

5.1.2.c. Gender

In order to analyse differences in scores between genders, an independent sample t-test was conducted to compare all **IMMERSION** factors between females ($n=159$) and males ($n=144$).

Table 29. Gender T-Tests for Factor Scores

Factor	Female	Male	T-statistic	df	p
R	4.53	4.46	0.42	301	0.67
I	4.24	3.92	2.31	301	0.02*
A	3.81	3.73	0.52	301	0.6
S	4.59	4.55	0.28	301	0.78
E	3.96	4.07	-0.68	301	0.49
C	3.77	3.89	-0.78	301	0.43

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

Table 30. Gender T-Tests for Sub-factors

Sub-factor	Female	Male	T-statistic	df	p
R1	5.14	4.92	1.05	301	0.29
R2	3.92	4	-0.36	301	0.72
I1	4.56	4.03	2.68	301	0.01*
I2	3.92	3.81	0.61	301	0.54
A1	3.71	3.5	1.08	301	0.28
A2	3.91	3.96	-0.28	301	0.78
S1	4.79	4.7	0.45	301	0.65
S2	4.4	4.4	-0.01	301	1
E1	4.32	4.53	-0.96	301	0.34
E2	3.59	3.61	-0.08	301	0.93
C1	3.95	4.02	-0.39	301	0.7
C2	3.59	3.76	-0.81	301	0.42

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

The t-tests found significant differences between males and females on the Investigative factor. Females were found to be scored significantly more (*female*=4.24, *male* = 3.92, $t(301)=2.31$, $p=.02$). However, this difference can only be seen in I1 (*female*=4.56, *male* = 4.03, $t(301)=2.68$, $p=.01$), which suggests that females may have a stronger inclination

toward seeking knowledge, reflection, and learning. Such gender differences in interest domains are consistent with broader patterns in vocational interest research, which often show that women gravitate more toward people-oriented, idea-oriented, or learning activities, whereas men tend to lean more toward technical domains.

6. Additional Information and Considerations

All assessments by *Central Test* have been developed for HR purposes only, and no clinical conclusions can be drawn from the results, not even from extreme scores. **IMMERSION** was developed for use in a range of organisational settings. It is important that the appropriate scores and interpretations are used in order to guide HR decisions for recruitment and selection, career selection and development. Accurate interpretation of psychometric tests helps both administrators and candidates to:

- put the assessment into context
- effectively compare candidates across target capabilities
- accurately identify areas for development
- maximise the fit between job requirements and candidate abilities and interests
- promote the organisation as a fair and objective place to work

Training and knowledge of the assessment process are crucial, as scores should be interpreted with considerations of reliability, validity, standard error of measurement, and any accommodations made during the assessment. Results should be presented within the context of the assessment and in a clear, meaningful and appropriate way for the target audience.

6.1. Interpreting Scores

All Central Test evaluations have been designed for use in HR contexts. Therefore, no clinical conclusions can be drawn from the results, even if they are extreme.

Reports provide three different levels of interpretive comments, determined by the user's score. The following table shows an example of how a specific level of one factor would be described. Extreme scores are defined as the lowest or highest possible values of a factor. In **IMMERSION**, the minimum score is 0 and the maximum is 10. Because of the ipsative format of the evaluation, it is impossible to obtain extreme scores across all factors simultaneously.

Each factor is composed of two subfactors. High scores in a given factor indicate a strong probability that the person has a genuine interest in that area. Mid-range scores (e.g., 4 to 6) indicate similarity with the normative group and may reflect a lower level of individual preference. However, all comments should be considered as hypotheses to be complemented with other sources of information, such as interviews, additional assessments, CV content, and personal references.

Table 31. *Example Descriptors of Different Levels of a Trait in Report Comments*

Level 1 (score 0 to 3)	Level 2 (score 3 to 7)	Level 3 (score 7 to 10)
Indicates a low interest	Indicates an interest	Indicates a strong interest

Combined profiles are calculated based on the two highest-scoring interests, which have similar importance, while the third factor has slightly less weight. For example, if a participant's combined profile is SRE (Social, Realistic, Enterprising), their main area of interest would be social activities (S) involving communication and working with others. They would also show a strong preference for outdoor or physical activities (R). Lastly, though to a lesser degree, they would enjoy a job where they can exert some influence over others and launch projects (E).

6.2. Report Content

The **IMMERSION** report is designed for career development and presents the ranking of a person's career interests, along with information on education and professional suitability. Below are its sections, along with useful details for interpretation.

Header: The header at the top of the report provides the assessment name, the date it was taken and the completion time. It allows for information verification and tracking.

Overview: This section introduces the **IMMERSION** assessment and its underlying model. It also outlines important considerations for interpreting the results accurately, ensuring that the report is understood within the right context and used appropriately.

Profile: This section provides a summary of the three dominant interest domains, derived from the participant's scenario-based responses, with a narrative description highlighting key tendencies and examples of suitable professional fields.

Interests Overview: This section highlights the highest and lowest areas of interest.

Results Table: The table summarises scores for the six RIASEC dimensions on a 0–10 scale, visually indicating the strength of preference for each factor. A check mark (✓) is positioned along the scale to show where the participant's score falls

Graph: The graph provides a quick and easy way to view the scores. It is a radar graph, with a highlighted dashed line indicating the midpoint. The score for each factor is shown in brackets. The purpose of the graph is to make interpretation easier by offering a visual summary of the profile.

Personalised comments by factor: This section provides the complete profile information.

- **Score:** Includes the score displayed on a scale from 0 to 10. In general, a score above 6.5 indicates a high level of interest, between 3.5 and 6 indicates a medium level, and below 3 indicates a low level.
- **Definition:** Includes the definition of the assessed factor, serving as a reference for interpreting the score. It is accompanied by information on potential skills, learning style, activity sector and examples of social and leisure activities related to that interest.

Combined profile and professional fit: Career interests often manifest in combinations, and analysing them together provides a better understanding of the profile. In the report, 15 combinations of career interests are ranked based on the scores obtained. For each combination, the percentage of matches is shown, with details on the relevant factors and ideal scores. A definition is also provided, along with examples of professions that best match the combination.

Matchings with studies and professions: This section lists professions based on the percentage of match and the results obtained in the assessment, with details on the relevant factors and ideal scores. A match by professional sector is also presented, including guidance on which studies or training programmes may be relevant.

6.3. Organisational Need: Career Selection, Recruitment and Development

IMMERSION is an effective tool for both the employer, for the selection of candidates on the basis of the occupational areas that suit their vocational interests, and the candidate, for the selection of their suitable career path based on their recreational interests. Choosing an appropriate career path is a difficult task, and with the number of working fields that exist, it can be confusing to know which one to take. By choosing a

career that links to their interests, the individual can find greater job satisfaction.

Recruitment and selection is an important organisational function which can be very time-consuming and expensive. In order to successfully hire and retain a candidate who will feel satisfied with their job, the job satisfaction of the employee becomes significant to the recruiter. By using **IMMERSION**, the recruiter will be able to infer how suitable a job may be for a candidate based on the assessment results. If the endorsed interests do not align with the job role, the job role may not be a good match. Bad recruitment decisions also have additional effects on organisational productivity, capability and sustainability. It is recommended that, prior to administration, recruiters have a detailed job description and person specifications based on a job analysis. Recruiters can also customise job profiles in their client accounts by specifying desired scores for each factor, which allows them to easily compare the target profile with a candidate's profile.

IMMERSION can also be used in line with performance evaluations, to help with discussing job performance in relation to occupational and vocational interests. Performance is key for employee development and success, and reviews should be a fair assessment of an employee's achievements and capabilities, giving the opportunity to discuss expectations and accomplishments. If an individual is not performing as they could be at work, it may be down to a lack of interest in a particular area of their job. Identifying this can help the recruiter to assess whether the employee should be transferred to another area of the company or to a different aspect of a project.

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

References

- Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta-analysis of the relationship between the five-factor model of personality and Holland's occupational types. *Personnel Psychology*, 56, 45– 74.
<http://dx.doi.org/10.1111/j.1744-6570.2003.tb00143.x>
- Brown, A., & Maydeu-Olivares, A. (2011, 06). Item Response Modeling of Forced-Choice Questionnaires. *Educational and Psychological Measurement*, 71(3), 460-502.
doi:10.1177/0013164410375112
- Bullock, E. E., Andrews, L. M., Braud, J., & Reardon, R. C. (2009-2010, Winter). Holland's theory in an international context: Applicability of RIASEC structure and assessments. *Career Planning and Adult Development Journal*, 25(4), 29-58.
- 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?* Germany: Lambert Academic Publishing.
- Cruise, P. A. (2012). The stability and organisational value of personality assessment. *Psyche Magazine Vol. 64, Spring*, pg 14-15. UK: Psychometrics Forum. Retrieved from
<http://www.psychometricsforum.org/uploads/Psyche/Psyche%2064%20-%20Spring%202012.pdf>
- Dingemanse, S. A., De Fruyt, F., Vergauwe, J., & Wille, B. (2015). A closer look at the psychological diversity within Holland interest types: Construct validation of the Career Insight. *Consulting Psychology Journal: Practice and Research*, 67(3), 234-257.
- Dingemanse, S. A., Van Amstel, B., De Fruyt, F., & Wille, B. (2007). *Loopbaan Inzicht Vragenlijst. Handleiding [Career Insight Questionnaire. Manual]*. Amsterdam: Harcourt Assessment.
- 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>

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>

Gottfredson, G. D. (1999). John L. Holland's contributions to vocational psychology: A review and evaluation. *Journal of Vocational Behaviour*, 55, 15-40.

Holland, J. L. (1958). A personality Inventory Employing Occupational Titles. *Journal of Applied Psychology*, 42(5), 336 – 342.

Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, 6(1), 35-45. <https://doi.org/10.1037/h0040767>

Holland, J. L. (1985). *Vocational preference inventory (VPI): Professional manual*. Odessa, Florida, US: Psychological Assessment Resources.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments (3rd ed.)*. Odessa, Florida, US: Psychological Assessment Resources.

Holland, J.L. & Holland, J. E. (1977). Vocational indecision: More evidence and speculation. *Journal of Counseling Psychology*, 24(5), 404-424.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Articles*, 2.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.

<http://doi.org/10.1080/10705519909540118>

Kaplan, R. M., & Saccuzzo, D. P. (2008). *Psychological Testing: Principles, Applications, and Issues (7th ed.)*. Belmont, CA: Wadsworth Publishing Co Inc.

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six Interests and Big Five Personality Factors. *Journal of Vocational Behavior*, 61, 217-239.

- Leutner, F., Codreanu, S.-C., Brink, S & Bitsakis, T. (2023). Game based assessments of cognitive ability in recruitment: Validity, fairness and test-taking experience. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.942662>
- Martin, B. A., Bowen, C.-C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality and Individual Differences*, 32(2), 247–256. [https://doi.org/10.1016/S0191-8869\(01\)00021-6](https://doi.org/10.1016/S0191-8869(01)00021-6)
- 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>
- Nauta, M. M. (2010). The Development, Evolution, and Status of Holland's Theory of Vocational Personalities Reflections and Future Directions for Counseling Psychology. *Journal of Counseling Psychology*, 11-22.
- 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>
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijsma. *Psychometrika*, 74(1), 145–154. <https://doi.org/10.1007/s11336-008-9102-z>
- 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)*.
- 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>

CONTACT US

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

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



www.centraltest.com

© 2025 by Central Test – Empowering Talent.
All rights reserved.