prospect screener

TECHNICAL /USER

MANUAL

DEVELOPED BY

JVR PSYCHOMETRICS

For more information, contact research@jvrafrica.co.za or any of the offices below.

First edition: Published 2013

Second edition: Published 2017

JVR Psychometrics forms part of **The JVR Africa Group**

JVR PSYCHOMETRICS (PTY) LTD

JOHANNESBURG OFFICE

15 Hunter Street | Ferndale | Randburg | 2194
P.O. Box 2560 | Pinegowrie | 2123 | Johannesburg | South Africa
T +27 11 781 3705/6/7 | F +27 11 781 3703 | E info@jvrafrica.co.za

CAPE TOWN OFFICE

3RD Floor | Gihon Building | Cnr of Bill Bezuidenhout Avenue &, Sportica Crescent
Tygervalley | Cape Town | 7530
P.O. Box 5534 | Tygervalley | 7536 | Cape Town | South Africa
T +27 861 333 916 | F +27 861 004 225 | E capeinfo@jvrafrica.co.za

FREE STATE OFFICE

8 Hugenote Street | Jordania | Bethlehem | 9700 | South Africa
T +27 58 303 3843 | E freestate-info@jvrafrica.co.za

PSYCHOMETRICS

ACKNOWLEDGEMENTS

The Prospect Screener is the result of a team effort, and many people have contributed to the success of the product over the years. The brainchild of the Centre of Research Excellence (CORE, previously known as the JvR Research department), the Prospect Screener was the first in the range of JVR-developed products. We would like to thank the following people and institutions for helping us to develop and improve the Prospect Screener continuously:

- Members of the Research team who contributed content to the Prospect Screener, namely Nicola Taylor, CJ van Zyl, and Jürgen Becker.
- Members of the Research team who spent time creating reports, manuals, conducting analysis, and creating norms, namely Brandon Morgan, Paul Vorster, and Mandy Wigdorowitz.
- JVR employees who were instrumental in the design and layout of reports and manuals, as well as proofreading content, namely Morné Stander, Alexis Kahts, and Elena Rust.
- Lizelle van Rooyen in the JVR IT department in her tireless commitment to get the Prospect Screener running on the JvR Online platform with the help of AddUp Holdings as developers.
- Harambee for the initial trial and feedback on the research version of the Prospect Screener.
- JVR's trusted clients for their continued use of the Prospect Screener.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABOUT THE AUTHOR

The Prospect Screener was designed and developed by the Centre of Research Excellence (CORE) department of JVR Psychometrics (hereafter referred to as JVR). JVR was established in 1993 with the focus of importing, researching, and distributing psychological tests. Today, JVR represents highly respected international test publishers and distributors for the Southern African region. JVR also actively sources, develops, and promotes psychological tests of high quality, developed in Southern Africa. JvR takes great pride in being South African, and is privileged to have shown significant growth over the past two decades. Distributing and publishing in excess of 200 tests, JVR is able to provide assessments in all the specialist fields of psychology including corporate, clinical, educational, neurological, counselling, and forensic use. Psychology as a science is dependent on continuous research, which is a cornerstone of JVR's values. CORE is committed to continuously evaluating the statistical integrity of the assessments developed and distributed by JVR.

1. INTRODUCTION

1.1 MAIN FEATURES/BRIEF OVERVIEW

The Prospect Screener is a screening tool used to measure basic verbal and numerical ability, detail-orientation, and workstyles (dependability and emotional stability). The purpose of this instrument is to screen candidates for further assessment, selection, and/or consideration rather than to predict or explain behaviour in its own right. It is thus used to determine if candidates are potentially good prospects for further psychometric assessment based on their inherent verbal and numerical ability and behavioural orientations. The Prospect Screener is an online assessment hosted on JvR Online and consists of 62 items that take approximately 15 to 30 minutes to complete.

1.2 PURPOSE AND RATIONALE

The purpose of the Prospect Screener is to screen out candidates that do not meet the minimum standards for employability or job performance based on their verbal and numerical ability, detail-orientation and workstyles. As psychological assessment processes are expensive, labour intensive, and time-consuming, assessing candidates who do not have the minimum characteristics or capacities to meet the job requirements reduces the efficiency and cost effectiveness of the assessment process. The Prospect Screener allows large numbers of prospective employees to be assessed both quickly and inexpensively. Candidates who are unlikely to perform effectively on more detailed, time-consuming, and expensive assessments are identified and screened out. Those candidates who meet minimum requirements are classified as prospects for further assessment and consideration.

1.3 ADMINISTRATION OVERVIEW/COMPONENTS AND TESTING PROCESS

The Prospect Screener is an online assessment tool where clients manage their own assessment profiles through a Prospect Screener assessment portal. The Prospect Screener is hosted on the JvR Online Portal where a number of assessment suites can be integrated. Credits for the use of the Prospect Screener may be purchased through JVR for a certain number of assessments. Alternatively, clients can purchase yearly licenses to assess larger groups of people.

The Prospect Screener is used for the corporate screening of applicants entering the organisation and **is not an educational, clinical, counselling, correctional, research or preventative assessment**. As the assessment is administered, scored, and interpreted through JvR Online, any human resource professional may use a report from the assessment for screening purposes. Please note that the assessment **is not used to predict ability or behavioural characteristics**. The assessment is only used to determine a candidate's potential suitability for further selection and consideration based on their ability and behavioural orientations. The Prospect Screener is recommended for jobs that require **entry-level** verbal ability, numerical ability, detail orientation, dependability, and emotional stability. The Prospect Screener may also be used for large talent pools where there is doubt that the group has the rudimentary capacities to meet certain job requirements.

2. THE CONCEPTUAL FRAMEWORK

2.1 HISTORY OF THE PRODUCT

The Prospect Screener was conceptualised and developed in 2012 and released in 2013 by JVR Psychometrics. The assessment was developed due to the extensive demand for a screening tool that can identify prospective candidates with the minimum capacities and characteristics to qualify for further testing and assessment. Many screening tools measure the basic skills, abilities, and knowledge of candidates to complete specific tasks or jobs effectively. However, no screening tool that measures basic psychological characteristics for screening purposes has been developed. The Prospect Screener fills this gap by providing a preliminary screening of verbal/numerical ability, detail-orientation, and workstyles (dependability and emotional stability). These attributes are widely considered to be important for effective job-fit and performance (Gatewood, Field, & Barrick, 2011; Hunter & Hunter, 1984; Muchinsky, Kriek, & Schreuder, 2004; Roth, Bobko, & McFarland, 2005).

2.2 THEORETICAL BACKGROUND/CONCEPTUAL FRAMEWORK

French (1974) and Beach (1970) describe the process of personnel selection as “sorting out or elimination of those judged unqualified to meet job and organization requirements” (Beach, 1970, p. 232). It can also be seen as “the process of choosing from a group of applicants the individual best suited for a particular position” (Grobler, Wärnich, Carrell, Elbert, & Hatfield, 2002, p. 174). In general the personal selection process has to determine from a large pool of applications those candidates that are best suited to meet the requirements of the job for which they are being assessed (Grobler et al., 2002). The Prospect Screener is particularly well suited for the initial

stage of the selection process where it is necessary to identify the most promising individuals for further screening (interviews, psychometric assessment, reference checking, etc.) and to determine those individuals who are potentially best suited to the job-tasks (Muchinsky et al., 2004).

The initial or preliminary screening phase usually consists of relatively cost effective and easy to implement techniques that are used to determine the suitability of a candidate for a particular job or position (van der Merwe, 2002). Screening techniques that are typically used in this stage include reference checking, prior training, job-experience, biographical blanks, job interviews, and résumé (curriculum vitae) checking (Gatewood et al., 2011), to name a few. While many of these techniques are valid methods for screening out candidates based on salient job requirements, they do not necessarily reveal psychological constructs that are assessed at a later phase in the selection process (Gatewood et al., 2011). They are also not considered strictly objective and standardised for all selection contexts (Muchinsky et al., 2004). The Prospect Screener is well suited to fill the need as an objective and standardised measure.

The Prospect Screener measures basic entry-level verbal and numerical ability, detail orientation, and workstyles that have been demonstrated, through extensive research, to be important predictors of job-performance. Research has shown, for example, that verbal and numerical ability, detail orientation, dependability, and emotional stability are important constructs for the prediction of job-performance, upward mobility, occupational attainment, job-satisfaction, positive/negative work perception, and supervisory style (Barrick & Mount, 1991; Chiswick & Miller, 2010; Kuncel, Hezlett, & Ones, 2004; Muchinsky et al., 2004; Rothmann & Coetzer, 2007; Roth et al., 2005; Salgado, 2003; Schmidt & Hunter, 2004; Spector, 1982).

3. SCALES OF THE PROSPECT SCREENER

The Prospect Screener measures verbal ability, numerical ability, detail orientation, and workstyles (dependability and emotional stability) through four scales. These scales are:

3.1 WORDS SCALE (8 ITEMS)

This scale measures entry-level verbal ability. Candidates are asked to match words of similar meaning. This scale measures the candidates' understanding of the meaning of English words. This ability is important for most jobs in South Africa where the *lingua franca* is English. Candidates with high scores on this scale can be considered to have minimum proficiency in using the English language. Candidates with low scores may find it difficult to understand and define English words and may not be considered proficient in the English language.

3.2 NUMBERS SCALE (10 ITEMS)

The Numbers scale measures entry-level numerical ability. Candidates are asked to complete mathematical problems that are composed of simple addition, subtraction, multiplication, and division. Some items in the Numbers scale also assess BODMAS (Brackets, Orders, Division, Multiplication, Addition, and Subtraction) which is akin to *mathematical grammar* required for most practical mathematical calculations. Candidates with high scores on this scale have the ability to do simple mathematical calculations, and are considered to have basic numeracy. Candidates who score low on this scale may be unable to do simple mathematical calculations, and may be considered innumerate.

3.3 DETAILS SCALE (24 ITEMS)

The Details scale measures detail orientation or detail-checking ability. Candidates are presented with words and numbers and are asked to determine whether the presented copies are identical. Detail orientation is important for most jobs requiring clerical work. High scores on the Details scale indicate a propensity to perceive differences in stimuli, which facilitates accuracy when completing detail-oriented work. Candidates with low scores on this scale may not be able to perceive differences in stimuli and may therefore make careless mistakes.

3.4 WORK STYLES SCALES (20 ITEMS)

The Work Styles scale measures a candidate's behavioural orientation towards work environments. The scale is composed of two dimensions: Dependability (10 items) and Emotional Stability (10 items). The scale measures whether candidates are dependable (conscientious, rule-abiding, careful, and detail conscious) and emotionally stable (not prone to anxiety, self-doubt, and vacillation of mood). The two dimensions of the Work Styles scale are important considerations for jobs that require emotionally stable, rule-abiding, risk-averse, and detail conscious workers.

4. INTERPRETING AND USE OF RESULTS

The Prospect Screener has four scales: Words, Numbers, Details, and Work Styles (Dependability and Emotional Stability). An Overall Prospect score is also reported. Each of the scale scores and Overall Prospect score are reported on a scale from 1 to 3. A score of 1 is a screened out result. This indicates that the candidate does not meet the minimum requirements on this scale and should probably not be considered for further assessment and selection. A score of 2 indicates that the candidate meets the minimum requirements on this scale and should be considered a prospect for further assessment and selection. A score of 3 indicates that the candidate is a good prospect and should definitely be considered for further assessment and selection. It is important to note that the Overall Prospect score should be consulted when making a decision about retaining or screening the candidate. The Overall Prospect score summarises the overall potential the candidate has for further assessment and selection and should be the first score consulted. The other scales of the Prospect Screener are used to contextualise results, and should not be interpreted independently.

The Prospect Screener Individual Report indicates whether or not a candidate should be screened out or if they are potentially a prospect for further assessment and selection. The report provides six scores: Words (basic verbal ability), Numbers (basic numerical ability), Details, Dependability, Emotional Stability, and Overall Prospect. The Overall Prospect score is a composite summary for the four scales of the Prospect Screener and is the only score that should be used when making pre-selection decisions. The Overall Prospect score is indicated by narrative descriptions of three screening levels: Screened-Out, Prospect, and Good Prospect.

5. DEVELOPMENT AND STANDARDISATION

The items for the Prospect Screener were selected from numerous item banks used in assessment and research. The items have been shown in previous research to have adequate psychometric properties. The items with the best psychometric properties were selected for inclusion in the Prospect Screener. The constructs assessed include verbal and numerical ability, detail orientation, dependability, and emotional stability. Research has shown that these constructs are most important for entry-level white collar positions (Roth et al., 2005; Chiswick & Miller, 2010; Kuncel et al., 2004; Muchinsky et al., 2004; Rothmann & Coetzer, 2007; Salgado, 2003; Schmidt & Hunter, 2004; Spector, 1982). The initial version of the Prospect Screener was composed of 87 items. After conducting research on the initial version of the Prospect Screener 25 items were removed. These items were removed to reduce the time a candidate will use to complete the instrument and to improve its psychometric properties. Only items with the best psychometric properties were retained for the current version of the Prospect Screener.

5.1 STANDARDISATION OF EACH SCALE

Verbal ability is measured by the Words scale. The items of the Words scale are composed of a standard word that candidates are asked to match to numerous alternatives (multiple choice format). The standard must be matched to a word option that most closely resembles the meaning of the standard. Each item in the Words scale has only one correct answer from an option of four alternative words (there is only one alternative word that most closely matches the meaning of the standard word). The Words scale therefore measures the candidate's ability to comprehend the meaning of words. This exercise determines whether a candidate is able to define

words used in everyday language; a requirement for most entry-level white collar positions. The Words scale was initially composed of 10 items, which was reduced to 8 items after initial research.

Numerical ability is measured through the Numbers scale. The items of this scale are characterised by numerical problems that make use of addition, subtraction, division, and multiplication to test for basic numeracy. Each item has a numerical problem that needs to be worked out and answered. There is only one correct option. The candidate has to select the correct answer from a series of alternative answers (four alternatives). The candidates **may not use a calculator** to answer the questions, but may use paper to make calculations by hand. The Numbers scale determines whether candidates have the capacity to do simple mathematical calculations. The exercise thus measures the functional numeracy (the ability to complete simple mathematical calculations) of a candidate; a requirement for most entry-level white collar positions. The initial version of the Numbers scale was composed of 10 items, all of which were retained after initial psychometric investigation.

Detail-orientation is measured through the Details scale. This scale is made up of verbal and numerical information that is replicated in a copy column. The candidate is asked to determine whether the copy column contains exactly the same information as the original column. Small differences exist between the information in some of the original and copy columns. The candidate chooses from two options, 'correct' or 'incorrect'. In the 'correct' option there are no differences in the information in the original and copy columns (the copy is correct). The 'incorrect' option is selected if there are differences between the information in the original and copy columns (the copy is incorrect). These items measure whether a candidate is detail conscious and able to differentiate between information with small (subtle) differences. This is an especially important construct for white collar workers who are required to detect

small mistakes or differences in written information. The scale was originally composed of 42 items, which were reduced to 24 items after initial psychometric investigation.

The Work Styles scale measures a candidate's personality-related characteristics of dependability (conscientiousness) and emotional stability (neuroticism). Conscientiousness measures a candidate's propensity for rule adherence, perfectionism, predictability, and risk-aversion. Neuroticism determines to what degree the candidate remains calm under pressure, how anxious the candidate is, and whether the candidate is likely or unlikely to vacillate between emotional extremes. High levels of dependability (high levels of conscientiousness) and emotional stability (low levels of neuroticism) are considered advantageous for most white collar jobs. The items on these scales ask the candidate to indicate on a five point Likert-type scale (Strongly Disagree, Disagree, Neither Agree or Disagree, Agree, or Strongly Agree) to what degree they agree or disagree with a statement which is characteristic of either dependability or emotional stability. The items are all positively keyed with no reverse scored items. Therefore, the higher candidate's scores on either of these scales are the more of the trait of dependability and emotional stability the candidate has. Each work style was composed of 10 items (20 items for both styles), which were all retained after initial psychometric investigation.

6. NORMATIVE SAMPLE

The normative sample consists of 10 422 respondents. There were more women (*n* = 6370, 61.1%) than men (*n* = 4049, 38.9%) in the sample. Three (.03%) respondents did not indicate their gender. The majority of respondents indicated that their home language was English (*n* = 2641, 25.3%). Other home languages spoken by the respondents were Zulu (*n* = 2258, 21.7%), Xhosa (*n* = 2241, 21.5%), Pedi (*n* = 684, 6.6%), Sotho (*n* = 648, 6.2%), Tswana (*n* = 630, 6.0%), Tsonga (*n* = 442, 4.2%), Afrikaans (*n* = 425, 4.1%), Venda (*n* = 262, 2.5%), Ndebele (*n* = 52, .5%), Swazi (*n* = 19, .2%), and Other (*n* = 62, .6%). Fifty-eight respondents (.6%) did not indicate their home language.

7. RESULTS

7.1 DESCRIPTIVE STATISTICS

Descriptive statistics for the Prospect Screener scale scores are presented in Table 1. Histograms and bean plots for the scale scores are provided in Figure 1 to Figure 5. Inspection of the figures indicates that the scale scores were mostly negatively skewed (i.e., most people scored high on the scales). This would be expected of a screening instrument where discrimination should occur at lower levels of the underlying latent trait.

Table 1: *Descriptive Statistics for the Prospect Screener Scale Scores*

	Mean	SD	Range	Median	Skewness	Kurtosis	SE
Words	6.63	1.36	8	7	-1.22	1.92	.01
Numbers	6.13	1.99	10	6	-.11	-.25	.02
Details	2.44	3.68	19	22	-1.43	1.30	.04
Dependability	43.62	4.87	40	44	-1.37	4.04	.05
Emotional Stability	37.49	6.23	40	38	-.46	.17	.06

Note. SE = Standard error of the mean.

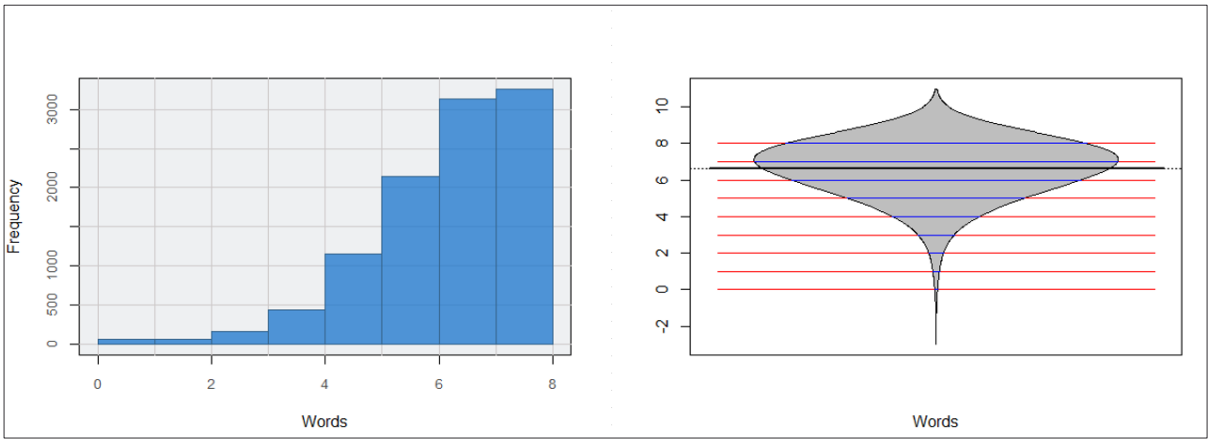

Figure 1: Histogram and bean plot for the Words scale

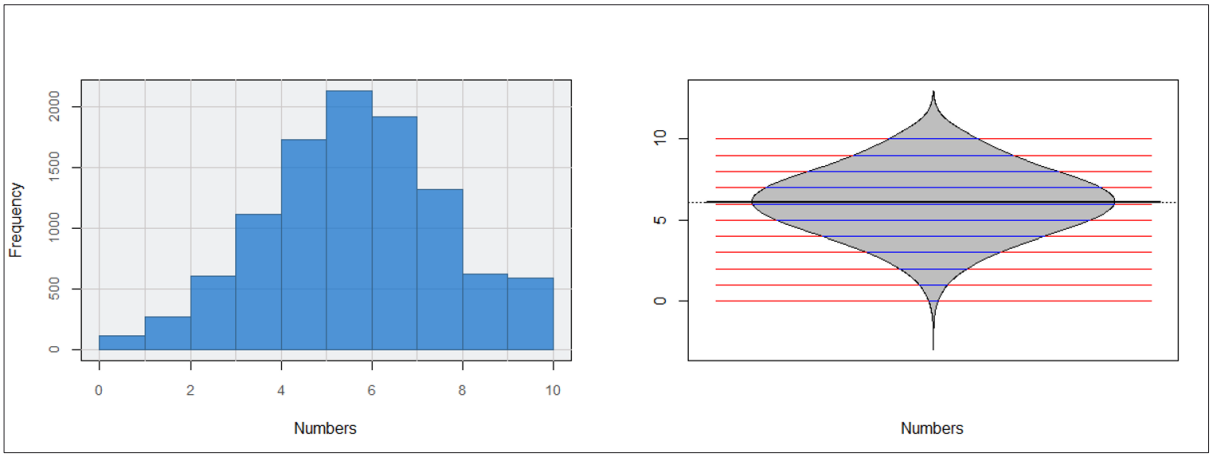

Figure 2: Histogram and bean plot for the Numbers scale

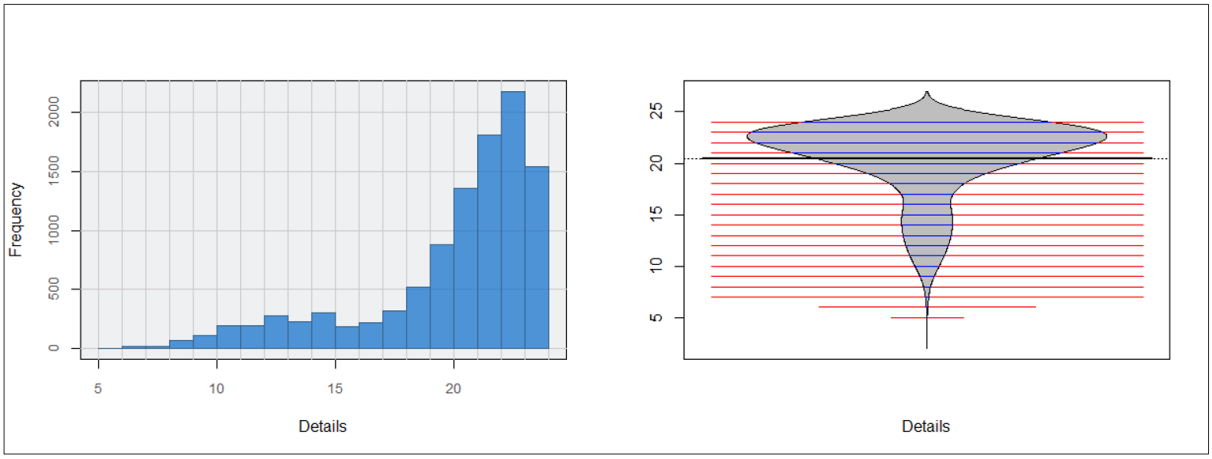

Figure 3: Histogram and bean plot for the Details scale

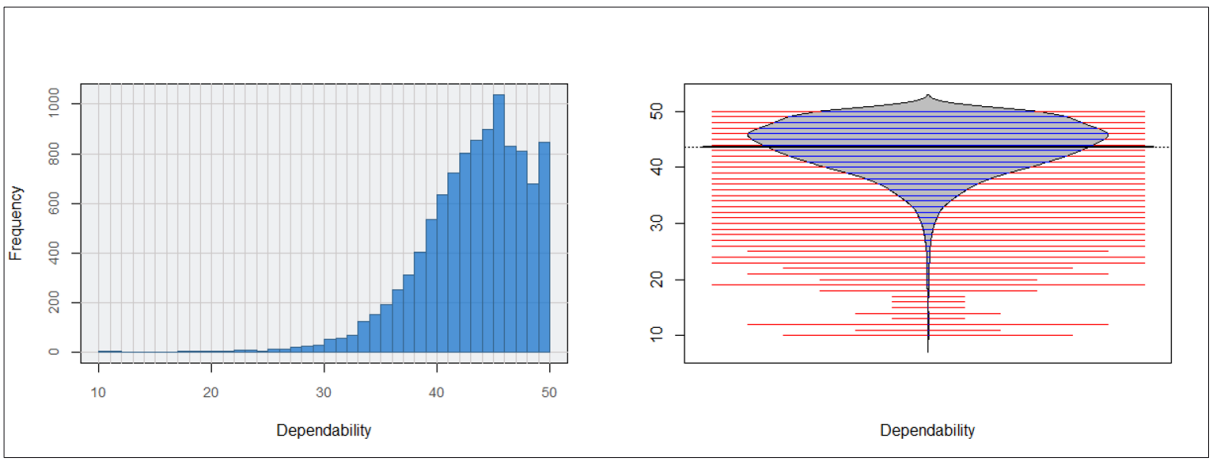

Figure 4: Histogram and bean plot for the Dependability scale

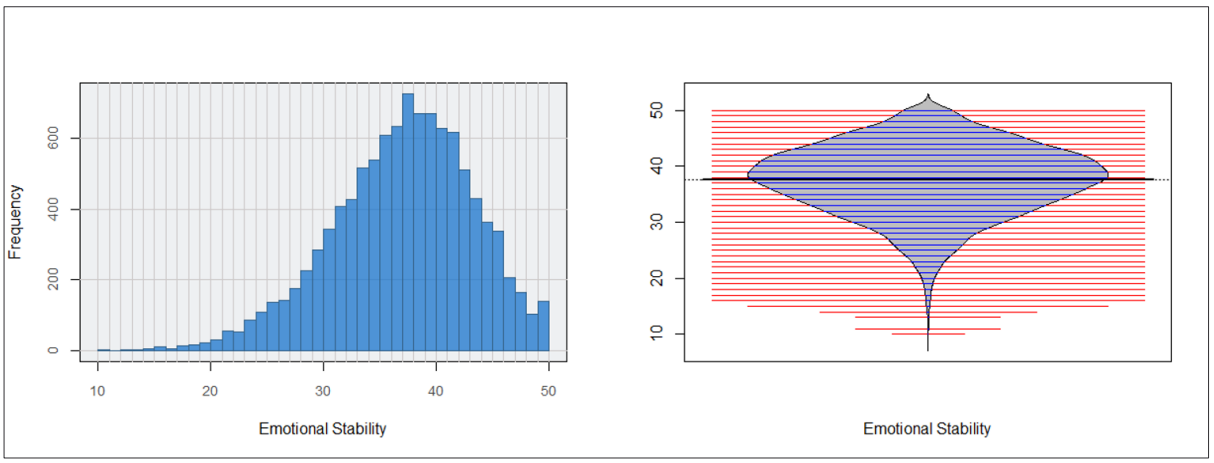

Figure 5: Histogram and bean plot for the Emotional Stability scale

7.2 SCALE RELIABILITIES

Reliability coefficients for the scale items of the Prospect Screener are presented in Table 2[1]. Cronbach alpha coefficients (Cronbach, 1951), Ordinal alpha coefficients

[1] Reliability coefficients on each scale for gender and language are presented in Appendix A.

(Zumbo, Gadermann, & Zeisser, 2007), Guttman's Lambda 2 (Guttman, 1945), and McDonald's Omega total (McDonald, 1999; Revelle & Zinbarg, 2009) were used. Item and person separation reliability estimates from the Rasch analysis are also presented. Ordinal alpha uses a tetrachoric/polychoric correlation matrix as input rather than a covariance matrix (Gadermann, Guhn, & Zumbo, 2012). Given the categorical nature of the data, Ordinal alpha may be warranted (Zumbo et al., 2007). Item and person reliability indices indicate the extent to which item/person locations can be reproduced (Linacre, 2016a). The reliability coefficients for the Words, Numbers, and Details section may be under-estimated because there is limited variance in the responses (i.e., most people correctly answered the items).

Table 2: *Reliability Coefficients for the Prospect Screener Scale Items*

Scale	α	O α	λ_2	ω	IR	PR
Words	.54 (.52 - .55)	.82	.55	.59	1.00	.08
Numbers	.63 (.62 - .65)	.79	.65	.64	1.00	.59
Details	.82 (.82 - .83)	.93	.83	.84	1.00	.51
Dependability	.73 (.72 - .74)	.83	.73	.75	1.00	.45
Emotional Stability	.72 (.71 - .73)	.77	.72	.72	1.00	.67

Note. 95% confidence intervals for α in parentheses, IR = Item separation reliability, PR = Person separation reliability.

7.3 ITEM DESCRIPTIVE STATISTICS

The average inter-item correlation coefficients and item-total correlation coefficients for the Prospect Screener scale items are presented in Table 3 to Table 7. The item-total correlation coefficient is an estimate of the correlation coefficient between the item and the total summated scale score. Positive item-total correlation coefficients indicate that the item is able to discriminate between people who score high and low

on the scale (Murphy & Davidshofer, 2005). For the Words, Numbers and Details scales, the mean item values can also be interpreted as a measure of item difficulty, where higher mean values indicate easier items[2] (Kerlinger & Lee, 2000). The average inter-item correlation coefficients and item-total correlation coefficients were positive across all the scales.

Table 3: *Item Level Descriptive Statistics for the Words Scale*

Item	Mean	SD	Average R	R Cor	R Drop
W1	.97	.16	.14	.45	.31
W2	.97	.16	.16	.32	.21
W3	.94	.24	.15	.36	.24
W4	.92	.27	.14	.42	.28
W5	.86	.35	.14	.45	.33
W6	.62	.49	.16	.30	.24
W7	.67	.47	.15	.33	.27
W8	.67	.47	.15	.39	.32

Note. Average R = average inter-item correlation, R Cor = item whole correlation corrected for item overlap and scale reliability, R Drop = item whole correlation against the scale without the item.

Table 4: *Item Level Descriptive Statistics for the Numbers Scale*

	Mean	SD	Average R	R Cor	R Drop
N1	.90	.30	.16	.22	.17
N2	.84	.37	.15	.35	.28
N3	.76	.42	.14	.40	.32
N4	.81	.39	.15	.28	.22
N5	.82	.38	.15	.36	.29
N6	.75	.43	.14	.44	.35
N7	.52	.50	.15	.35	.29
N8	.31	.46	.14	.39	.32
N9	.21	.41	.14	.51	.35
N10	.22	.41	.14	.55	.39

Note. Average R = average inter-item correlation, R Cor = item whole correlation corrected for item overlap and scale reliability, R Drop = item whole correlation against the scale without the item.

[2] Under classical test theory these item difficulties are sample dependent. It is therefore better to investigate the ranking of the item locations (difficulties) under the Rasch model.

Table 5: *Item Level Descriptive Statistics for the Details Scale*

	Mean	SD	Average R	R Cor	R Drop
D1	.75	.43	.18	.45	.43
D2	.87	.34	.19	.19	.16
D3	.94	.25	.18	.36	.29
D4	.79	.41	.18	.31	.30
D5	.85	.35	.18	.28	.22
D6	.96	.20	.18	.31	.26
D7	.73	.44	.18	.45	.44
D8	.90	.31	.17	.56	.50
D9	.85	.36	.17	.56	.53
D10	.89	.32	.17	.57	.50
D11	.82	.38	.17	.47	.45
D12	.92	.27	.17	.54	.47
D13	.72	.45	.18	.41	.40
D14	.90	.30	.17	.59	.51
D15	.68	.47	.18	.36	.35
D16	.90	.31	.17	.61	.54
D17	.65	.48	.18	.36	.36
D18	.93	.26	.17	.53	.47
D19	.96	.20	.18	.40	.33
D20	.93	.26	.18	.39	.32
D21	.94	.23	.17	.49	.39
D22	.77	.42	.18	.33	.31
D23	.95	.21	.18	.33	.26
D24	.86	.34	.18	.30	.29

Note. Average R = average inter-item correlation, R Cor = item whole correlation corrected for item overlap and scale reliability, R Drop = item whole correlation against the scale without the item.

Table 6: *Item Level Descriptive Statistics for the Dependability Dimension*

	Mean	SD	Average R	R Cor	R Drop
DP1	3.91	1.14	.20	.37	.32
DP4	4.23	1.10	.20	.46	.40
DP6	4.04	1.14	.20	.48	.41
DP7	4.47	.97	.22	.54	.46
DP8	4.06	1.14	.20	.54	.45
DP11	2.98	1.26	.20	.51	.43
DP14	3.67	1.27	.20	.53	.45
DP16	3.82	1.18	.19	.55	.47
DP18	3.27	1.23	.20	.24	.20
DP20	3.05	1.28	.20	.48	.40

Note. Average R = average inter-item correlation, R Cor = item whole correlation corrected for item overlap and scale reliability, R Drop = item whole correlation against the scale without the item.

Table 7: *Item Level Descriptive Statistics for the Emotional Stability Dimension*

	Mean	SD	Average R	R Cor	R Drop
ES2	3.91	1.14	.20	.47	.39
ES3	4.23	1.10	.20	.48	.40
ES5	4.04	1.14	.20	.42	.35
ES9	4.47	.96	.22	.24	.21
ES10	4.06	1.14	.20	.50	.42
ES12	2.98	1.25	.20	.42	.36
ES13	3.67	1.27	.20	.42	.36
ES15	3.82	1.18	.19	.56	.48
ES17	3.27	1.23	.20	.47	.40
ES19	3.05	1.28	.20	.47	.40

Note. Average R = average inter-item correlation, R Cor = item whole correlation corrected for item overlap and scale reliability, R Drop = item whole correlation against the scale without the item.

7.4 RASCH ANALYSIS

A Rasch[3] analysis was conducted on each scale of the Prospect Screener[4]. Items with Infit and Outfit mean squares values (IMNSQ and OMNSQ) ≥ 1.40 were considered to be underfitting items, and items with IMNSQ and OMNSQ values ≤ .60 to be overfitting items (Bond & Fox, 2007; Wright, Linacre, Gustafon, & Martin-Lof, 1994). Underfitting items are particularly problematic because they degrade the quality of the scale (Bond & Fox, 2007). IMNSQ investigates unexpected responses on items that are targeted at the respondent's underlying latent ability measure, whereas OMNSQ investigates unexpected responses to items that are either too easy or too difficult for the respondent (Linacre, 2016c). It may be more appropriate to investigate IMNSQ item fit statistics because OMNSQ is sensitive to outliers (Bond & Fox, 2007).

7.4.1 WORDS SCALE

Item locations and fit statistics for the Words scale are provided in Table 8. The item locations ranged from -2.08 to 1.78 logits. One item (W1) demonstrated overfit based on the Outfit MNSQ. No items demonstrated underfit. The test information function is provided in Figure 6. The test information function is the sum of the item information functions at a given ability level. Information is defined as the reciprocal of the standard error and indicates the precision with which a person's ability can be estimated at different ability estimates (Baker, 2001). The test information function indicates that most of the information is below the 0 logit ability level. This is what would be expected from a screening tool.

[3] The rating scale model was used for polytomous items.

[4] The Winsteps software (Linacre, 2016b) was used for all Rasch analyses.

Table 8: *Words Scale Item Location and Item Fit Statistics*

			Infit		Outfit		PT-Measure	
Item	Location	SE	MNSQ	Z	MNSQ	Z	Corr	Exp
W1	-2.08	.07	.86	-2.8	**.56**	-5.2	.37	.31
W2	-2.14	.07	1.03	.5	1.05	.5	.30	.30
W3	-1.06	.05	1.00	.0	1.15	2.4	.37	.39
W4	-.72	.04	.97	-1.0	.98	-.4	.42	.41
W5	.19	.03	.93	-3.9	.87	-4.5	.52	.49
W6	2.20	.03	1.09	8.8	1.20	9.5	.59	.63
W7	1.82	.03	1.04	3.5	1.11	6.0	.59	.61
W8	1.78	.03	.95	-5.3	1.03	1.8	.62	.60
Mean	.00	.04	.98	.0	.99	1.2		
SD	1.66	.02	.07	4.2	.19	4.6		

Note. OMNSQ ≥ 1.4 or ≤ .60 in bold.

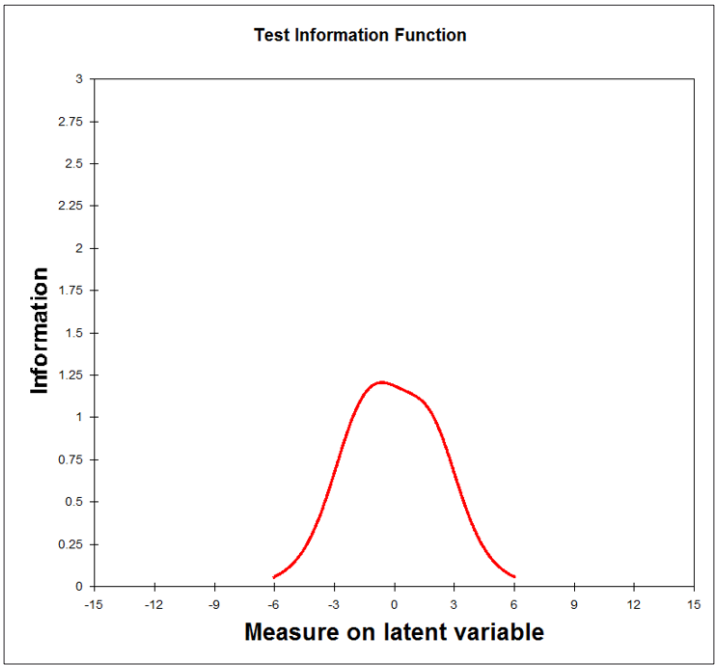

Figure 6: Test information function for the Words scale

7.4.2 NUMBERS SCALE

Item locations and fit statistics for the Numbers scale are provided in Table 9. The item locations ranged from -2.17 to 2.74 logits. No items demonstrated overfit. Two items demonstrated underfit (N1 and N4) on the Outfit MNSQ but not the Infit MNSQ. The test information function is provided in Figure 7. The test information function indicated that there was more information at lower levels of ability. This is again what is expected from a screening tool.

Table 9: *Numbers Scale Item Location and Item Fit Statistics*

			Infit		Outfit		PT-Measure	
Item	Location	SE	MNSQ	Z	MNSQ	Z	Corr	Exp
N1	-2.17	.04	1.04	1.7	**1.69**	9.3	.30	.34
N2	-1.52	.03	.96	-2.5	1.13	3.0	.41	.39
N3	-.91	.03	.94	-4.7	.97	-.8	.47	.45
N4	-1.25	.03	1.04	2.8	**1.47**	9.9	.38	.42
N5	-1.36	.03	.95	-2.9	1.13	3.1	.42	.41
N6	-.82	.03	.91	-7.3	.84	-5.7	.50	.45
N7	.64	.02	1.05	5.0	1.16	8.7	.51	.54
N8	1.96	.03	1.05	3.7	1.17	5.7	.55	.57
N9	2.74	.03	.93	-4.2	1.38	7.7	.58	.57
N10	2.71	.03	.88	-7.1	1.11	2.5	.61	.57
Mean	.00	.03	.98	-1.5	1.21	4.3		
SD	1.76	.00	.06	4.3	.23	4.7		

Note. OMNSQ ≥ 1.4 or ≤ .60 in bold.

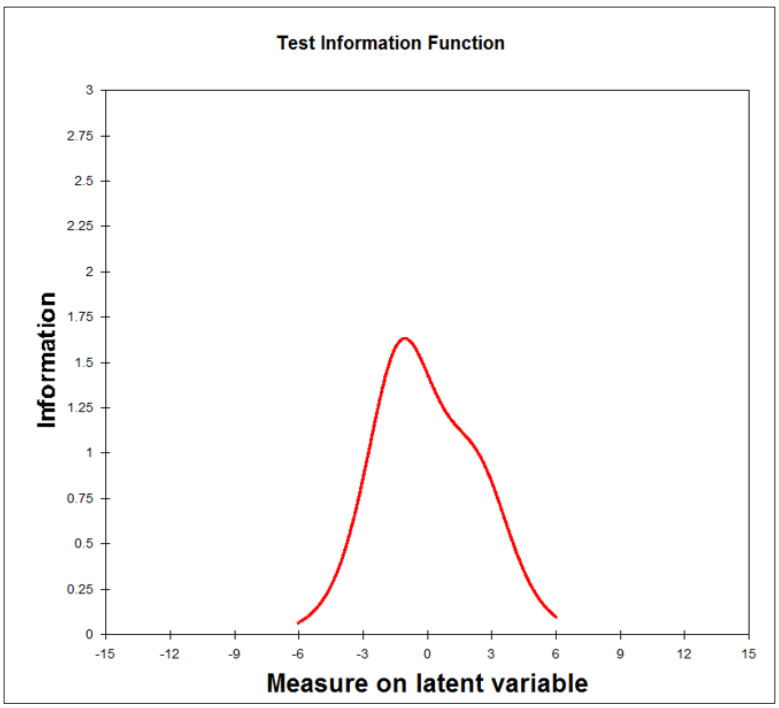

Figure 7: Test information function for the Numbers scale

7.4.3 DETAILS SCALE

Item locations and fit statistics for the Details scale are provided in Table 10. The item locations ranged from -1.51 to 1.75 logits. Three items demonstrated overfit (D16, D19, and D21) on the Outfit MNSQ. Two items demonstrated underfit (D2 and D5) on the Outfit MNSQ but not the Infit MNSQ. The test information function is provided in Figure 8. Most of the test information is located at the 0 logit ability.

Table 10: *Details Scale Item Location and Item Fit Statistics*

			Infit		Outfit		PT-Measure	
Item	Location	SE	MNSQ	Z	MNSQ	Z	Corr	Exp
D1	1.10	.00	1.00	.3	1.01	.6	.51	.51
D2	.03	.96	1.26	9.9	**1.49**	9.9	.27	.42
D3	-.94	.00	1.01	.4	.95	-.7	.31	.31
D4	.80	.96	1.16	9.9	1.13	5.3	.41	.49
D5	.19	.00	1.22	9.9	**1.51**	9.9	.30	.44
D6	-1.47	.96	.98	-.4	1.09	1.1	.26	.26
D7	1.22	.00	1.00	.0	1.00	.1	.52	.52
D8	-.30	.96	.84	-7.4	.69	-7.7	.47	.39
D9	.23	.00	.84	-9.1	.71	-9.7	.53	.44
D10	-.19	.96	.85	-7.3	.75	-6.5	.48	.40
D11	.49	.00	.96	-2.5	.88	-4.3	.49	.46
D12	-.63	.96	.85	-5.9	.63	-8.0	.44	.35
D13	1.30	.00	1.04	3.0	1.04	2.4	.50	.52
D14	-.34	.96	.83	-8.0	.67	-8.2	.48	.38
D15	1.59	.00	1.11	9.0	1.13	8.2	.48	.54
D16	-.32	.96	.80	-9.3	**.60**	-9.9	.50	.38
D17	1.75	.00	1.07	6.3	1.10	6.2	.51	.55
D18	-.82	.96	.84	-5.8	.56	-8.9	.43	.33
D19	-1.51	.00	.94	-1.5	**.60**	-5.6	.31	.25
D20	-.76	.96	1.00	-.1	1.01	.2	.33	.33
D21	-1.12	.00	.92	-2.6	**.54**	-8.1	.36	.30
D22	.95	.96	1.15	9.6	1.12	5.4	.43	.50
D23	-1.33	.00	1.01	.2	.85	-2.0	.28	.27
D24	.09	.96	1.13	6.3	1.10	2.7	.37	.43
Mean	.00	.04	.99	.2	.92	-1.1		
SD	.96	.01	.13	6.4	.27	6.3		

Note. OMNSQ ≥ 1.4 or ≤ .60 in bold.

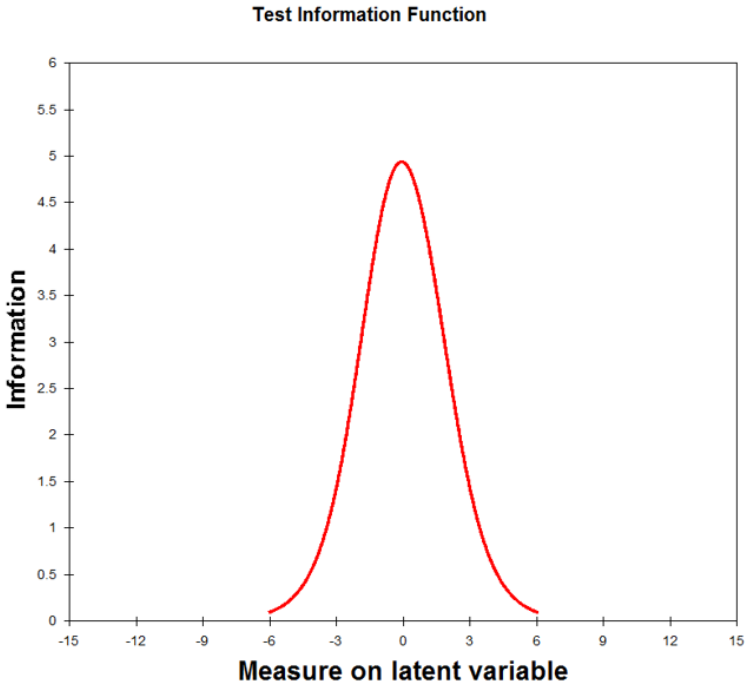

Figure 8: Test information function for the Details scale

7.4.4 DEPENDABILITY SCALE

Item locations and fit statistics for the Dependability scale are provided in Table 11. The item locations ranged from -.83 to 1.60 logits. No items demonstrated overfit or underfit. The test information function is provided in Figure 9. Most of the test information is located at the -1 to 1 logit range.

Table 11: *Dependability Scale Item Location and Item Fit Statistics*

			Infit		Outfit		PT-Measure	
Item	Location	SE	MNSQ	Z	MNSQ	Z	Corr	Exp
DP1	-.33	.01	1.07	4.2	1.10	5.4	.39	.39
DP4	-.47	.01	1.17	9.2	1.19	9.4	.39	.37
DP6	-.74	.01	1.19	9.4	1.14	6.4	.40	.34
DP7	-.25	.01	.78	-9.9	.83	-9.9	.45	.40
DP8	-.64	.01	.95	-2.8	.91	-4.6	.46	.36
DP11	-.41	.01	.95	-3.0	.98	-1.3	.41	.38
DP15	1.60	.01	1.09	7.0	1.27	9.9	.38	.49
DP17	1.12	.01	1.03	2.9	1.12	9.1	.39	.50
DP19	.95	.01	1.04	3.2	1.09	7.1	.44	.49
DP20	-.83	.01	.97	-1.3	.89	-5.4	.44	.33
Mean	.00	.01	1.02	1.9	1.05	2.6		
SD	.83	.00	.11	5.8	.13	6.9		

Note. OMNSQ ≥ 1.4 or ≤ .60 in bold.

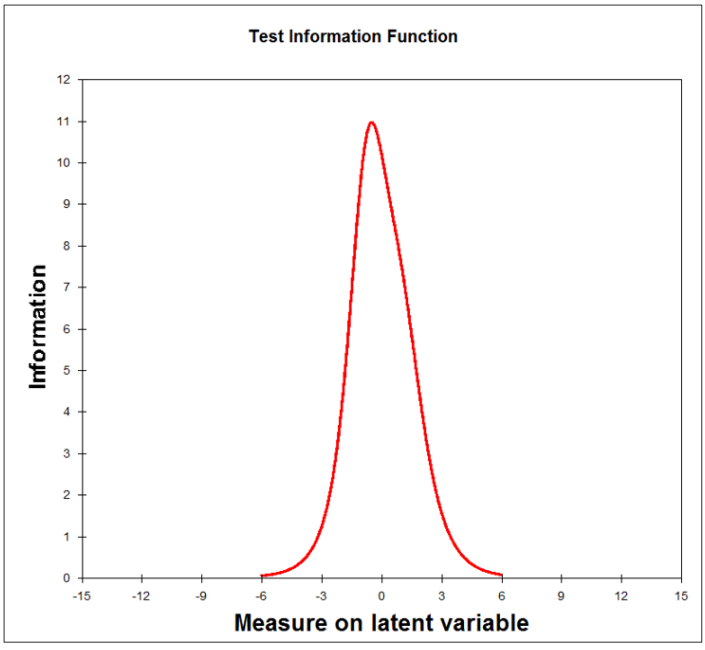

Figure 9: Test information function for the Dependability scale

7.4.5 EMOTIONAL STABILITY SCALE

Item locations and fit statistics for the Emotional Stability scale are provided in Table 12. The item locations ranged from -.85 to .73 logits. No item demonstrated overfit. One item demonstrated underfit (ES9) on both the Outfit and Infit MNSQ. The test information function is provided in Figure 1. As with the Dependability scale, most of the test information is located at approximately -1 to 1 logits of ability.

Table 12: *Dependability Scale Item Location and Item Fit Statistics*

			Infit		Outfit		PT-Measure	
Item	Location	SE	MNSQ	Z	MNSQ	Z	Corr	Exp
ES2	-.11	.01	.96	-3.2	.98	-1.5	.49	.49
ES3	-.49	.01	1.11	6.9	1.05	2.8	.47	.44
ES5	-.26	.01	1.08	5.7	1.09	5.0	.46	.47
ES9	-.85	.01	**1.43**	9.9	**1.49**	9.9	.31	.39
ES10	-.28	.01	1.01	.8	.99	-.4	.51	.47
ES12	.73	.01	.98	-1.9	1.00	.1	.53	.56
ES13	.12	.01	1.09	7.1	1.12	7.8	.49	.51
ES15	-.02	.01	.86	-9.9	.85	-9.9	.56	.50
ES17	.48	.01	.90	-8.8	.92	-5.9	.54	.55
ES19	.67	.01	.94	-4.7	.97	-2.7	.56	.56
Mean	.00	.01	1.04	.2	1.05	.5		
SD	.48	.00	.15	6.6	.17	5.8		

Note. OMNSQ ≥ 1.4 or ≤ .60 in bold.

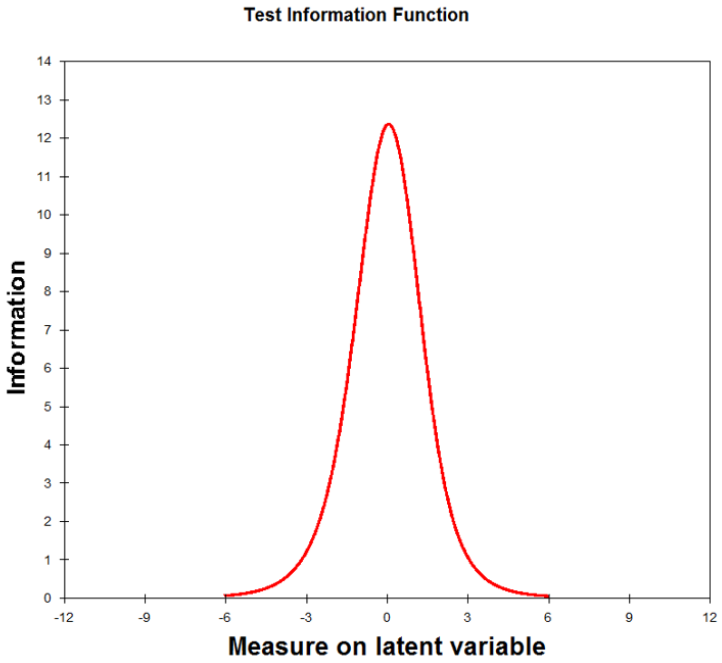

Figure 10: Test information function for the Emotional Stability scale

7.5 CONSTRUCT VALIDITY

Construct validity for each scale was investigated using confirmatory factor analysis. The weighted mean and variance adjusted estimator (WLSMV; Muthén & Muthén, 1998) with delta parametrisation was used for each analysis because of the categorical nature of the indicator variables (Brown, 2015).

7.5.1 WORDS SCALE

The fit statistics for the Words scale were: WLSMV χ^2 (20) = 211.697, p < .001, CFI = .965, TLI = .951, and RMSEA = .030 (.027 - .034). However, the baseline RMSEA was < .158 and therefore the CFI and TLI may not be informative. The standardised factor loadings ranged from .430 to .840. The average variance extracted (AVE) was .379.

Table 13: *Standardised Loadings for the Words Scale*

Item	Estimate	SE	Z-Value	*p*
W1	.84	.02	35.83	< .001
W2	.63	.03	21.64	< .001
W3	.59	.02	25.76	< .001
W4	.64	.02	33.58	< .001
W5	.66	.02	4.28	< .001
W6	.43	.02	26.83	< .001
W7	.48	.02	3.27	< .001
W8	.58	.02	36.06	< .001

7.5.2 NUMBERS SCALE

The fit statistics for the Numbers scale were: WLSMV χ^2 (35) = 2323.445, $p < .001$, CFI = .910, TLI = .884, RMSEA = .079 (.076 - .082). Inspection of the modification indices suggested that item error terms for N9 and N10 ($\delta_{N9,N10}$ =1799.076, EPC = .1.021) be allowed to covary. These error terms were therefore allowed to covary because both of these items measured BODMAS (i.e., the error terms would be expected to correlate). The fit statistics for the revised model were: WLSMV χ^2 (34) = 288.031, $p <$ .001, CFI = .990, TLI = .987 RMSEA = .027 (.024 - .030). The standardised factor loadings ranged from .375 to .643. The AVE was .269.

Table 14: *Standardised Loadings for the Numbers Scale*

Item	Estimate	Std.Err	Z-Value	*p*
N1	.38	.02	19.15	< .001
N2	.54	.02	34.32	< .001
N3	.59	.01	42.18	< .001
N4	.42	.02	25.62	< .001
N5	.55	.02	35.48	< .001
N6	.64	.01	48.41	< .001
N7	.49	.01	34.56	< .001
N8	.59	.02	37.71	< .001
N9	.43	.02	23.99	< .001
N10	.50	.02	29.17	< .001

7.5.3 DETAILS SCALE

The fit statistics for the Details scale were: WLSMV χ^2 (252) = 6968.015, p < .001, CFI = .906, TLI = .898, RMSEA = .051 (.050 - .052). The standardised factor loadings ranged from .273 to .883. The AVE was .403.

Table 15: *Standardised Loadings for the Details Scale*

Item	Estimate	Std.Err	Z-Value	*p*
D1	.62	.01	52.17	< .001
D2	.27	.02	15.98	< .001
D3	.57	.02	35.78	< .001
D4	.43	.01	3.37	< .001
D5	.41	.02	25.74	< .001
D6	.55	.02	28.82	< .001
D7	.62	.01	54.09	< .001
D8	.79	.01	73.92	< .001
D9	.79	.01	76.86	< .001
D10	.80	.01	78.09	< .001
D11	.66	.01	55.44	< .001
D12	.78	.01	7.30	< .001
D13	.59	.01	48.77	< .001
D14	.81	.01	8.53	< .001
D15	.52	.01	42.56	< .001
D16	.83	.01	86.39	< .001
D17	.52	.01	42.38	< .001
D18	.79	.01	7.85	< .001
D19	.67	.01	47.09	< .001
D20	.60	.02	39.59	< .001
D21	.74	.01	62.40	< .001
D22	.44	.01	31.26	< .001
D23	.55	.02	32.91	< .001
D24	.44	.02	28.85	< .001

7.5.4 DEPENDABILITY SCALE

The fit statistics for the Dependability dimension were: WLSMV χ^2 (35) = 599.129, p < .001, CFI = .977, TLI = .971, RMSEA = .039 (.037 - .042). The standardised factor loadings ranged from .329 to .659. The AVE was .33.

Table 16: *Standardised Loadings for the Dependability Scale*

Item	Estimate	Std.Err	Z-Value	*p*
DP1	.46	.01	43.96	< .001
DP4	.56	.01	59.07	< .001
DP6	.60	.01	61.23	< .001
DP7	.60	.01	74.80	< .001
DP8	.63	.01	71.38	< .001
DP11	.60	.01	69.66	< .001
DP14	.66	.01	77.12	< .001
DP16	.66	.01	75.00	< .001
DP18	.33	.01	3.90	< .001
DP20	.59	.01	58.97	< .001

7.5.5 EMOTIONAL STABILITY SCALE

Table 17: *Standardised Loadings for the Emotional Stability Scale*

Item	Estimate	Std.Err	Z-Value	*p*
ES2	.52	.01	56.60	< .001
ES3	.57	.01	58.98	< .001
ES5	.48	.01	48.81	< .001
ES9	.31	.01	23.81	< .001
ES10	.58	.01	64.26	< .001
ES12	.45	.01	49.25	< .001
ES13	.46	.01	47.25	< .001
ES15	.63	.01	77.21	< .001
ES17	.50	.01	56.59	< .001
ES19	.52	.01	58.34	< .001

The fit statistics for the Emotional Stability scale were: WLSMV χ^2 (35) = 1231.471, p < .001, CFI = .947 TLI = .932, RMSEA = .057 (.055 - .060). The standardised factor loadings ranged from .314 to .632. The AVE was .26.

7.6 DIFFERENTIAL ITEM FUNCTIONING

Uniform differential item functioning (DIF) was conducted using the Mantel-Haenszel (1959) and Mantel (1963) test within the Rasch model framework to determine if item locations were invariant across gender and language (Engelhard, 2013). The Mantel-Haenszel and Mantel tests within the Rasch framework are calculated on ability strata rather than raw score strata (Linacre, 2016a). The Mantel-Haenszel (1959) log-odds estimator and Liu-Agresti (1996) cumulative log-odds estimators are also provided. Classification values for these two log-odds estimators are: < .43 = negligible DIF, .43 to .64 = moderate DIF, and > .64 = large DIF (Penfield & Algina, 2006). Due to the large sample size the *p* values must be interpreted with caution.

7.6.1 WORDS SCALE

DIF across gender and language is presented in Table 18 to Table 21[5]. Inspection of Table 18 indicates that only one item demonstrated large DIF. The remainder of the items had negligible DIF.

[5] Bonferroni corrections were not applied to the *p* values for each item. The hypothesis relating to each item is therefore whether or not there is DIF on each item rather than DIF for the set of items (Linacre, 2016d).

Table 18: *DIF Across Women and Men on the Words Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
W1	-.70	31.54	.00	**-.84**
W2	.04	.04	.85	.04
W3	.08	1.21	.27	.12
W4	.04	.22	.64	.05
W5	-.15	6.85	.01	-.19
W6	.30	31.13	.00	.29
W7	-.40	48.86	.00	-.38
W8	.25	19.06	.00	.25

Note. CUMLOR = Cumulative log-odds ratio in logits.

Inspection of Tables 19 to 21 indicates that there were four items with large DIF across the English and Afrikaans groups. Six of the items demonstrated large DIF across the Afrikaans and African language groups. No items had large DIF across the English and African language group.

Table 19: *DIF Across English and Afrikaans on the Words Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
W1	.47	3.71	.05	**.71**
W2	-.18	.12	.73	-.14
W3	-.28	1.49	.22	-.29
W4	-.32	2.38	.12	-.32
W5	.43	6.91	.01	.55
W6	-1.54	85.30	.00	**-1.42**
W7	.69	19.49	.00	**.67**
W8	.67	18.81	.00	**.68**

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 20: *DIF Across English and African Languages on the Words Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
W1	.09	.16	.69	.08
W2	.15	.75	.39	.15
W3	.02	.03	.87	.02
W4	.43	19.85	.00	.43
W5	-.30	16.71	.00	-.34
W6	.17	7.15	.01	.16
W7	-.04	.65	.42	-.05
W8	-.14	5.16	.02	-.15

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 21: *DIF Across African Languages and Afrikaans on the Words Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
W1	.38	5.22	.02	**.78**
W2	-.33	1.42	.23	-.34
W3	-.30	1.98	.16	-.30
W4	**-.75**	17.59	.00	**-.79**
W5	**.73**	2.81	.00	**.89**
W6	**-1.71**	10.00	.00	**-1.55**
W7	**.74**	24.53	.00	**.71**
W8	**.82**	34.45	.00	**.88**

Note. CUMLOR = Cumulative log-odds ratio in logits.

7.6.2 NUMBERS SCALE

DIF across gender and language is presented in Tables 22 to 25. No items demonstrated large DIF across gender.

Table 22: *DIF Across Women and Men on the Numbers Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
N1	-.26	1.22	.00	-.23
N2	.06	1.33	.25	.08
N3	-.06	1.01	.31	-.06
N4	-.19	7.72	.01	-.16
N5	-.08	1.52	.22	-.08
N6	.21	14.77	.00	.23
N7	.08	2.86	.09	.08
N8	.16	8.24	.00	.16
N9	-.11	4.43	.04	-.14
N10	.00	.51	.47	-.05

Note. CUMLOR = Cumulative log-odds ratio in logits.

Inspection of Tables 23 to 25 indicates that there was no large DIF across any of the groups. One item for the Afrikaans and African languages demonstrated large moderate DIF.

Table 23: *DIF Across English and Afrikaans on the Numbers Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
N1	.10	.08	.77	.08
N2	-.06	.12	.73	-.08
N3	.24	2.14	.14	.25
N4	-.25	2.42	.12	-.24
N5	.04	.00	.96	.02
N6	.23	1.84	.18	.23
N7	.42	9.74	.00	.41
N8	-.19	2.89	.09	-.26
N9	-.41	3.90	.05	-.38
N10	-.33	2.52	.11	-.30

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 24: *DIF Across English and African Language Group on the Numbers Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
N1	.03	.14	.71	.03
N2	.06	.81	.37	.07
N3	.05	1.22	.27	.07
N4	-.02	.08	.77	-.02
N5	.00	.03	.85	-.02
N6	.03	.57	.45	.05
N7	-.22	16.20	.00	-.22
N8	-.06	1.49	.22	-.08
N9	.15	3.68	.06	.15
N10	.16	5.05	.03	.18

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 25: *DIF Across African Language Group and Afrikaans on the Numbers Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
N1	.07	.17	.68	.10
N2	-.13	.71	.40	-.16
N3	.19	1.04	.31	.17
N4	-.23	1.83	.18	-.20
N5	.04	.02	.89	.03
N6	.20	1.19	.28	.18
N7	.65	25.97	.00	.64
N8	-.13	1.30	.25	-.17
N9	-.55	9.53	.00	-.56
N10	-.48	7.99	.01	-.49

Note. CUMLOR = Cumulative log-odds ratio in logits.

7.6.3 DETAILS SCALE

DIF across gender and language is presented in Tables 26 to 29. Inspection of Table 26 indicates that no items displayed large DIF across gender.

Table 26: *DIF Across Women and Men on the Details Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
D1	-.12	4.20	.04	-.11
D2	.00	.12	.73	-.02
D3	.00	.01	.94	.01
D4	-.12	7.41	.01	-.15
D5	.17	3.40	.07	.11
D6	.00	.00	.99	.00
D7	-.10	3.28	.07	-.10
D8	-.15	2.40	.12	-.13
D9	-.02	.00	1.00	.00
D10	-.08	.25	.62	-.04
D11	.00	.21	.65	-.03
D12	.00	.72	.40	.08
D13	.20	12.57	.00	.19
D14	-.11	.89	.35	-.08
D15	.07	1.34	.25	.06
D16	-.31	15.97	.00	-.34
D17	.15	6.45	.01	.13
D18	-.29	9.93	.00	-.30
D19	-.23	3.11	.08	-.20
D20	.15	2.73	.10	.14
D21	-.05	.00	.99	.00
D22	.10	.92	.34	.05
D23	.03	.11	.74	.04
D24	.23	7.57	.01	.18

Note. CUMLOR = Cumulative log-odds ratio in logits.

Inspection of Tables 27 to 29 indicates that three items demonstrated large DIF across the three language groups. DIF was present for the English and Afrikaans language groups and the African and Afrikaans language groups. No statistically and practically significant DIF was present between the English and African Language groups.

Table 27: *DIF Across English and Afrikaans on the Numbers Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
D1	-.08	.13	.72	-.06
D2	-.05	.00	.95	.00
D3	.09	.01	.91	.06
D4	-.05	.03	.87	-.03
D5	-.36	3.54	.06	-.30
D6	-.40	2.00	.16	-.39
D7	.03	.03	.87	.03
D8	.17	.30	.59	.14
D9	-.46	9.60	.00	-.53
D10	.21	.85	.36	.24
D11	-.52	11.86	.00	-.55
D12	.54	4.32	.04	.58
D13	.45	7.82	.01	.45
D14	.41	3.07	.08	.48
D15	-.27	3.64	.06	-.25
D16	.49	5.10	.02	.54
D17	.72	23.10	.00	**.73**
D18	.06	.00	.98	.03
D19	.01	.00	.98	-.05
D20	-.09	.14	.71	-.11
D21	-.19	.65	.42	-.24
D22	-.11	.22	.64	-.08
D23	-.53	3.41	.07	-.50
D24	-.14	.20	.66	-.10

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 28: *DIF Across English and African Language Group on the Details Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
D1	.10	2.95	.09	.11
D2	.00	.19	.67	.03
D3	.02	.05	.83	.03
D4	.00	.00	.98	.00
D5	-.08	.51	.48	-.05
D6	-.09	.78	.38	-.12
D7	.00	.00	.96	.01
D8	.00	.06	.81	-.03
D9	.11	1.89	.17	.11
D10	.15	2.92	.09	.15
D11	-.06	.76	.39	-.06
D12	.06	.03	.87	.02
D13	-.03	.27	.61	-.03
D14	.08	.70	.40	.08
D15	.00	.01	.94	.01
D16	.17	3.64	.06	.18
D17	-.06	.68	.41	-.05
D18	.03	.01	.92	.02
D19	.13	.68	.41	.11
D20	.00	.00	.98	.00
D21	-.07	1.31	.25	-.13
D22	.00	.28	.60	.03
D23	-.26	4.74	.03	-.26
D24	-.25	7.70	.01	-.21

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 29: *DIF Across African Language Group and Afrikaans on the Details Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
D1	-.19	1.61	.20	-.19
D2	-.05	.03	.86	-.04
D3	.06	.01	.94	.05
D4	-.05	.06	.81	-.04
D5	-.29	2.27	.13	-.23
D6	-.30	1.17	.28	-.29
D7	.03	.04	.84	.04
D8	.17	.66	.42	.19
D9	-.57	17.17	.00	**-.67**
D10	.06	.07	.79	.08
D11	-.46	1.38	.00	-.48
D12	.48	4.35	.04	.55
D13	.48	1.10	.00	.49
D14	.33	2.42	.12	.41
D15	-.27	3.94	.05	-.24
D16	.31	2.50	.11	.37
D17	.78	28.63	.00	**.77**
D18	.03	.01	.91	.00
D19	-.12	.08	.77	-.12
D20	-.09	.16	.69	-.12
D21	-.12	.17	.68	-.13
D22	-.11	.64	.42	-.12
D23	-.28	1.17	.28	-.29
D24	.11	.29	.59	.11

Note. CUMLOR = Cumulative log-odds ratio in logits.

7.6.4 DEPENDABILITY SCALE

DIF across gender and language is presented in Tables 30 to 33. Inspection of these Tables indicates that no items displayed large DIF across gender. One item displayed large DIF across the Afrikaans and African language groups. The same item also displayed large moderate DIF across the English and Afrikaans language groups.

Table 30: *DIF Across Women and Men on the Dependability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
DP1	.00	.75	.39	.04
DP4	.11	19.16	.00	.20
DP6	.00	.49	.49	.03
DP7	.05	5.77	.02	.10
DP8	.11	24.07	.00	.23
DP11	.22	87.14	.00	.41
DP15	-.10	34.96	.00	-.23
DP17	.02	1.02	.31	.04
DP19	-.24	10.00	.00	-.51
DP20	.09	17.45	.00	.20

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 31: *DIF Across English and Afrikaans on the Dependability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
DP1	.31	21.91	.00	.58
DP4	.08	.75	.39	.10
DP6	.09	2.29	.13	.20
DP7	.15	9.94	.00	.34
DP8	.03	.23	.63	.06
DP11	-.02	.08	.79	-.03
DP15	-.03	1.07	.30	-.10
DP17	.04	.51	.48	.07
DP19	-.29	39.70	.00	**-.64**
DP20	-.13	3.01	.08	-.20

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 32: *DIF Across English and African Language Group on the Dependability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
DP1	.00	.00	.96	.00
DP4	.00	.06	.81	-.01
DP6	.00	.04	.84	.01
DP7	.00	.02	.90	.01
DP8	.06	5.88	.02	.12
DP11	.02	.85	.36	.04
DP15	-.03	2.07	.15	-.06
DP17	-.04	7.23	.01	-.11
DP19	.00	.21	.65	.02
DP20	.09	5.99	.01	.13

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 33: *DIF Across African Language Group and Afrikaans on the Dependability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
DP1	.31	25.14	.00	.58
DP4	.08	2.38	.12	.18
DP6	.09	1.93	.17	.18
DP7	.15	9.40	.00	.31
DP8	-.03	.16	.69	-.04
DP11	-.04	.54	.46	-.08
DP15	.00	.26	.61	-.05
DP17	.08	3.45	.06	.17
DP19	-.29	49.21	.00	**-.67**
DP20	-.21	8.32	.00	-.32

Note. CUMLOR = Cumulative log-odds ratio in logits.

7.6.5 EMOTIONAL STABILITY SCALE

DIF across gender and language is presented in Tables 34 to 37. Inspection of Table 34 indicates that no items displayed large DIF across gender. For language, one item

displayed large DIF for the Afrikaans language group in comparison to the English and African language groups.

Table 34: *DIF Across Women and Men on the Emotional Stability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
ES2	-.07	12.04	.00	-.14
ES3	.05	.87	.35	.04
ES5	.02	3.23	.07	.07
ES9	-.22	41.75	.00	-.31
ES10	.00	.19	.67	-.02
ES12	.07	26.68	.00	.20
ES13	-.16	58.59	.00	-.30
ES15	.06	5.58	.02	.10
ES17	-.09	21.51	.00	-.18
ES19	.20	10.00	.00	.45

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 35: *DIF Across English and Afrikaans on the Emotional Stability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
ES2	-.18	12.34	.00	-.37
ES3	-.13	3.94	.05	-.22
ES5	-.03	.22	.64	-.05
ES9	-.37	29.96	.00	-.61
ES10	.19	11.05	.00	.40
ES12	.05	1.19	.28	.11
ES13	-.09	2.12	.15	-.15
ES15	.07	2.40	.12	.16
ES17	-.03	.46	.50	-.07
ES19	.35	49.59	.00	**.75**

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 36: *DIF Across English and African Language Group on the Emotional Stability Scale*

Item	DIF Contrast	χ^2	*p*	CUMLOR
ES2	.06	8.71	.00	.13
ES3	-.03	2.97	.09	-.09
ES5	.03	1.99	.16	.07
ES9	.02	.44	.51	.04
ES10	-.07	9.91	.00	-.15
ES12	-.03	3.04	.08	-.07
ES13	.00	.99	.32	.04
ES15	.00	.04	.85	.01
ES17	.02	3.54	.06	.08
ES19	-.02	2.34	.13	-.07

Note. CUMLOR = Cumulative log-odds ratio in logits.

Table 37: *DIF Across African Language Group and Afrikaans on the Dependability Scale African and Afrikaans*

Item	DIF Contrast	χ^2	*p*	CUMLOR
ES2	-.24	24.64	.00	-.50
ES3	-.10	1.83	.18	-.14
ES5	-.05	1.28	.26	-.12
ES9	-.39	33.50	.00	-.62
ES10	.27	22.02	.00	.54
ES12	.08	3.74	.05	.18
ES13	-.09	3.75	.05	-.19
ES15	.07	1.94	.16	.14
ES17	-.06	2.65	.10	-.15
ES19	.38	64.09	.00	**.80**

Note. CUMLOR = Cumulative log-odds ratio in logits.

8. DIFFERENTIAL TEST FUNCTIONING

The combined effect of DIF across each scale was investigated using the differential test functioning procedures described by Penfield and Algina (2006). Their approach is an unsigned variance-based technique and is appropriate for dichotomous and polytomous items (Penfield, 2007; Penfield & Algina, 2006). It is based on the Mantel-Haenszel (1959) common odds ratio for dichotomous items and the Liu-Agresti (1996) cumulative common odds ratio for polytomous items. The variance of the generalised DIF effect is given as an unweighted and weighted value denoted by τ^2 for dichotomous items and v^2 for polytomous items. Weighted τ^2 and v^2 were used as indicators of differential test functioning in this analysis. Suggested differential test functioning interpretation criteria for τ^2 and v^2 are small = < .07, medium = .07 to .14, and large > .14 (Penfield & Algina, 2006). The DIFAS (Penfield, 2005) software was used to calculate v^2. The results are presented in Table 38. Inspection of Table 38 indicates that there was no differential test functioning across men and women and across English and African language respondents. For the Afrikaans language respondents differential test functioning was present for the Words scale. The Dependability and Emotional Stability scales also demonstrated some differential test functioning between the Afrikaans and African language respondents.

Table 38: *Differential Test Functioning*

Scale	Gender	Eng. and Afr.	Eng. and African	African and Afr.
Words	.06	.43	.02	.61
Numbers	.03	.05	.01	.10
Details	.01	.08	.00	.09
Dependability	.01	.07	.01	.12
Emo. Stability	.05	.10	.01	.14

Note. Emo. Stability = Emotional Stability, Eng = English, Afr = Afrikaans.

9. CORRELATION COEFFICIENTS

Pearson correlation coefficients and Spearman-rho rank order correlation coefficients for the Prospect Screener scales are reported in Table 39. Inspection of the non-parametric Loess regression lines (Cleveland, 1979) indicated that for the most part the relationships between the variables were linear. Inspection of multivariate normality using Mardia's coefficient (Mardia, 1970) and contour plots found that bivariate normality was not met across most of the variables. The correlation coefficients between the scales had small to medium effect sizes (Cohen, 1988).

Table 39: *Pearson and Spearman-Rho Rank Order Correlations for the Prospect Screener Scales*

	W	N	D	DP	ES
W	.	.29***	.26***	.03**	.19***
N	.28***	.	.35***	.02	.18***
D	.26***	.33***	.	.05***	.16***
DP	.06***	.03**	.07***	.	.25***
ES	.19***	.18***	.15***	.23***	.

Note. Pearson correlations below the diagonal, Spearman rho rank-order correlations above the diagonal. W = Words, N = Numbers, D = Details, DP = Dependability, ES = Emotional Stability scale. ** = $p < .01$, *** = $p < .001$.

10. GENDER DIFFERENCES IN MEAN SCORES

Differences in group centroids were investigated using Hoteling's T^2 test. The results indicate that there was a statistically significant difference in the group centroids between men and women [$T^2(5, 10413) = 52.024$, $p < .001$]. Post-hoc independent samples *t* tests with a Holm-Bonferroni correction were subsequently applied. The results indicate that there was a statistically significant difference in the mean scores for men and women on the Words[6] [*M* men = 6.79, SD = 1.31, *M* women = 6.53, SD = 1.37, $t(10300) = -9.314$, $p_{adj} < .001$, $d = .19$], and Numbers scales [*M* men = 6.35, SD = 2.09, *M* women = 5.98, SD = 1.90, $t(10300) = -9.208$, $p_{adj} < .001$, $d = .19$], and on the Emotional Stability dimension [*M* men = 38.38, SD = 5.99, *M* women = 36.92, SD = 6.33, $t(10300) = -11.636$, $p_{adj} < .001$, $d = .24$]. There were no statistically significant differences on the Details scale [*M* men = 2.40, SD = 3.78, *M* women = 2.46, SD = 3.62, $t(10300) = .796$, $p_{adj} = .464$, $d = .02$] or on the Dependability dimension [*M* men = 43.69, SD = 4.87, *M* women = 43.57, SD = 4.88, $t(10300) = -1.195$, $p_{adj} = .464$, $d = .02$]. The effect sizes for the statistically significant differences were all small (Cohen, 1988) indicating that these differences are mostly negligible.

[6] The required condition for *t* tests was not met for some of the scales (normality). The Wilcoxon Rank Sum Test was therefore performed to verify the results (this test investigates differences in locations rather than means). Because the sample size is large, and the Wilcoxon Rank Sum Test produced similar results, only the *t* test results are presented.

11. LANGUAGE DIFFERENCES IN MEAN SCORES

A one-way between groups MANOVA was used to investigate mean differences across the five scales for home language. The results indicated that there was a statistically significant difference between the three language groups on the combined dependent variables [V = .008, $F(10,20592) = 8.386$, $p < .000$, $\eta_p^2 = .004$]. The MANOVA analysis was followed up with one way ANOVAs[7]. Applying the Holm-Bonferroni correction found that there was only a statistically significant difference on the Words scale [$F(2,10299) = 23.419$, $p_{adj} < .001$, $\eta^2 = .005$]. Post-hoc comparisons using Tukey's HSD indicated that the mean Words score for the Afrikaans group [$M = 6.20$, SD = 1.67] was lower than the mean Words score for the English [$M = 6.62$, SD = 1.38, $p < .001$] and African language groups [$M = 6.66$, SD = 1.32, $p < .001$]. There was no statistically significant difference on the Numerical [$F(2, 10299) = 3.151$, $p_{adj} = .140$, $\eta^2 = .001$] and Details [$F(2, 10299) = 3.351$, $p_{adj} = .140$, $\eta^2 = .001$] scales or on the Dependability [$F(2, 10299) = 1.828$, $p_{adj} = .322$, $\eta^2 = .000$] and Emotional Stability [$F(2, 10299) = .457$, $p_{adj} = .633$, $\eta^2 = .000$] dimensions.

[7] Similar results were obtained using the Kruskal-Wallis test. These results are therefore not reported.

12. CONCLUDING COMMENTS

The psychometric properties of the Prospect Screener indicate that the assessment can be used effectively to assess the entry level capacities and characteristics of candidates. The assessment is based on the most prevalent competencies required by most jobs and can be used to successfully pre-screen large groups of candidates.

13. R PACKAGES

The following R (R Core Team, 2015) packages were used in the analysis:

Psych (Revelle, 2015), lessR (Gerbing, 2015), beanplot (Kampstra, 2008), outliers (Komsta, 2011), mvoutlier (Filzmoser & Gschwandtner, 2015), MVN (Korkmaz, Goksuluk, & Zararsiz, 2015), Lambda4 (Hunt, 2013), lavaan (Rosseel, 2012), semTools (semTools contributors, 2016), and DescTools (Signorell et al., 2016), as well as all associated dependencies.

REFERENCES

Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). United States of America: ERIC Clearing House on Assessment and Evaluation.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1–26.

Beach, D. S. (1970). *Personnel: The management of people at work*. New York, NY: Collier-Macmillan.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, NJ: Lawrence Erlbaum & Associates.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.

Chiswick, B. R., & Miller, P. W. (2010). Occupational language requirements and the value of English in the US labour market. *Journal of Population Economics*, *23*, 353–372. doi: 1.1007/s00148-008-0230-7

Cleveland, S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association, 74*(368), 829–836.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). United States of America: Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Engelhard, G., Jr. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences.* New York, NY: Routledge.

Filzmoser, P., & Gschwandtner, M. (2015). mvoutlier: Multivariate outlier detection based *on robust methods* (Version = 2..6). Retrieved from http://www.statistik.tuwien.ac.at/public/filz/

French, W. (1974). *The personnel management process*. Boston, MA: Houghton Mifflin.

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation, 17*(3), 1-13.

Gatewood, R. D., Field, H. S., & Barrick, M. (2011). *Human resource selection* (7th ed.). Australia: Thomson South-Western.

Gerbing, D. W. (2015). lessR: Less Code, More Results (R package version 3.3.1. Retrieved from http://cran.r-project.org/web/packages/lessR/index.html.

Grobler, P. A., Wärnich, S., Carrell, M. R., Elbert, N. F., & Hatfield, R. D. (2002). *Human resource management in South Africa* (2nd ed.). London, England: Thompson Learning.

Guttman, L. (1945). A basis for analysing test-retest reliability. *Psychometrika, 10*, 255–282.

Hunt, T. (2013). Lambda4: Collection of Internal Consistency Reliability Coefficients (Version 3.0). Retrieved from https://cran.r-project.org/web/packages/Lambda4/index.html

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*, 72–98.

Kampstra, P. (2008). Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software, 28*, 1-9.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of Behavioral Research*. Wadsworth/ Thomson Learning: Northridge, CA.

Komsta, L. (2011). outliers: Tests for outliers (Version .14.). Retrieved from http://www.r-project.org/

Korkmaz, S., Goksuluk, D., Zararsiz, G. (2014). MVN: An R Package for assessing multivariate normality. *The R Journal, 6*(2), 151-162.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job-performance: Can one construct/predict them all? *Journal of Personality and Social Psychology*, *86*(1), 148–161.

Linacre, J. M. (2016a). Reliability and separation of measures. Retrieved from http://winsteps.com/winman/reliability.htm

Linacre, J. M. (2016b). Winsteps® (Version 3.92.0) [Computer Software]. Beaverton, Oregon: Winsteps.com. Retrieved January 1, 2016. Available from http://www.winsteps.com/

Linacre, J. M. (2016c). Misfit diagnosis: Infit outfit mean-square standardized. Retrieved from http://www.winsteps.com/winman/diagnosingmisfit.htm

Linacre, J. M. (2016d). Bonferroni – multiple t-tests. Retrieved from http://www.winsteps.com/winman/bonferroni.htm

Liu, I. M., & Agresti, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics, 52*, 1223-1234.

Mantel N. (1963). Chi-square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association, 58*, 690-70.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies. *Journal of the National Cancer Institute*, *22*(4), 719-748.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519–553.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.

Muchinsky, P. M., Kriek, H. J., & Schreuder, A. M. G. (2004). *Personnel Psychology* (2nd ed.). Cape Town, South Africa: Oxford University Press.

Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (6th ed.). Upper Saddle River, NJ: Pearson Education.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. Los Angeles, CA: Authors.

Penfield, R. D. (2005). DIFAS: Differential item functioning analysis system. *Applied Psychological Measurement, 29*, 150-151.

Penfield, R. D. (2007). An approach for categorizing DIF in polytomous items. *Applied Measurement in Education, 20*(3), 335-355.

Penfield, R. D., & Algina, J. (2006). A generalized DIF effect variance estimator for measuring unsigned differential test functioning in mixed format tests. *Journal of Educational Measurement, 43*(4), 295-312.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Revelle, W. (2015). Psych: Procedures for Personality and Psychological Research (Version = 1.5.6.). Retrieved from http://CRAN.R-project.org/package=psych.

Roth, P. L., Bobko, P., & McFarland, L. A. (2005). A meta-analysis of work sample test validity: Integrating and updating some classic literature. *Personnel Psychology*, *58*, 1009–1037.

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega and the glb: Comments on Sijtsma. *Psychometrika, 74*(1), 145–154.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. Retrieved from http://www.jstatsoft.org/v48/i02/

Rothmann, S., & Coetzer, E. P. (2007). The big five personality dimensions and job performance. *South African Journal of Industrial Psychology*, *29*(1), 68–74.

Salgado, J. F. (2003). Predicting job-performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, *76*(3), 323–346.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job-performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173.

semTools Contributors. (2016). semTools: Useful tools for structural equation modeling. R package version .4-11. Retrieved from http://cran.r-project.org/package=semTools

Signorell, A., et. al. (2016). DescTools: Tools for descriptive statistics. R package version .99.16.

Spector, P. E. (1982). Behavior in organizations as a function of employee's locus of control. *Psychological Bulletin*, *91*(3), 482–497.

Van der Merwe, R. P. (2002). Psychometric testing and human resource management. *South African Journal of Industrial Psychology*, *28*(2), 77–86.

Wright, B. D., Linacre, J. M., Gustafson, J. E., & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*(3), 37.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21-29.

APPENDIX A: RELIABILITY COEFFICIENTS FOR GENDER AND LANGUAGE

Table 40: *Reliability Coefficients for Women*

Scale	α	O α	λ2	ω
Words	.53 (.51 - .55)	.81	.54	.58
Numbers	.60 (.58 - .61)	.76	.61	.60
Details	.82 (.81 - .83)	.93	.83	.84
Dependability	.72 (.71 - .74)	.82	.73	.74
Emotional Stability	.72 (.71 - .73)	.77	.73	.72

Table 41: *Reliability Coefficients for Men*

Scale	α	O α	λ2	ω
Words	.55 (.52 - .57)	.83	.56	.60
Numbers	.68 (.66 - .70)	.82	.69	.68
Details	.83 (.82 - .84)	.93	.84	.85
Dependability	.73 (.72 - .75)	.83	.74	.76
Emotional Stability	.71 (.69 - .72)	.76	.71	.71

Table 42: *Reliability Coefficients for English Speakers*

Scale	α	O α	λ2	ω
Words	.55 (.52 - .58)	.83	.57	.61
Numbers	.64 (.61 - .66)	.79	.65	.64
Details	.83 (.82 - .84)	.93	.84	.84
Dependability	.74 (.72 - .76)	.83	.74	.75
Emotional Stability	.74 (.72 - .76)	.79	.74	.74

Table 43: *Reliability Coefficients for Afrikaans Speakers*

Scale	α	O α	λ2	ω
Words	.68 (.62 - .75)	.87	.69	.71
Numbers	.61 (.55 - .68)	.78	.63	.62
Details	.83 (.80 - .86)	.93	.84	.84
Dependability	.76 (.71 - .81)	.85	.76	.76
Emotional Stability	.78 (.74 - .83)	.83	.79	.78

Table 44: *Reliability Coefficients for African Language Speakers*

Scale	α	O α	λ2	ω
Words	.51 (.49 - .53)	.81	.53	.57
Numbers	.63 (.62 - .65)	.79	.65	.64
Details	.82 (.81 - .83)	.93	.83	.84
Dependability	.72 (.71 - .73)	.82	.72	.74
Emotional Stability	.71 (.69 - .72)	.76	.71	.71