Mental Agility Series

South African English Manual Supplement

ebilities

Mental Agility Series Tests

Distributed in South Africa by JvR Psychometrics (Pty) Ltd.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

The Mental Agility Series (MAS) range of assessment batteries measures various cognitive abilities that are required to perform mental tasks. A person's innate cognitive ability provides the basis for determining their capacity to learn new skills, adapt to new situations, and develop a body of expert knowledge, as well as acts as a determinant of speed and accuracy on routine tasks. The assessments are grounded in psychometric theory and are based on individual differences in theory and research, providing results that are appropriate for a diverse range of assessment purposes. An in-depth discussion of the supporting theory and perspectives of measurement, along with the rationale behind the development of the assessment can be found in the international *ebilities GMA Series Technical Monograph* (Lewis & Cadman, 2017) and *Mental Agility Series Technical Manual* (Douglas, 2017).

1.1 User qualifications

According to the Health Professions Act No. 56 of 1974, measures of cognitive functioning are considered psychological assessments. By this definition, only trained psychology professionals who are registered with the Health Professions Council of South Africa may gain access to use the MAS range in South Africa.

No accreditation training is required for the registered professionals in South Africa to use the MAS range of assessments, however, JvR Psychometrics will gladly assist with demonstrating how the results of the assessment may be interpreted and applied, if required.

1.2 Appropriate use

The MAS range of assessment batteries are intended to be used with adults, not children or adolescents, and are intended for a normal population, not clinical, psychiatric, or psychopathological samples. Although most widely used in occupational contexts for personnel selection and development, the MAS range may also be appropriate for use by adults in personal development and research initiatives.

The MAS range are neither medical examinations, nor can they be used to evaluate medical conditions, mental illness, mental disabilities, or physical disabilities. In addition, inappropriate assessment uses include forecasting or evaluating neuropsychological behaviour, suicidal thoughts/behaviour, specific criminal actions, dementia, visual/motor abilities, hyperactivity, perceptual abilities, and/or information obtained from polygraph/biofeedback instruments.

To ensure that respondents understand the content of the items, it is recommended that the MAS be used with adults who have at least a Grade 12 (NQF Level 4) level of education and a corresponding level of English literacy.

2 Mental Agility Series Tests

The MAS range of assessment batteries consists of combinations of the following primary tests:

- **Financial Reasoning –** a measure of quantitative knowledge; items involve quantitative manipulation of numbers based on financial concepts such as interest rates and taxable income.

- **Numerical Operations –** a measure of quantitative knowledge; items test basic mathematical and computational skill involving addition, subtraction, division, and multiplication.

- **Series –** a measure of inductive reasoning; items test ability to determine the next number in a series, based on an induced logical rule.

- **Swaps –** a measure of fluid ability and working memory; items test ability to determine correct sequence of pictures after swapping their order.

- **Vocabulary –** a measure of crystallised ability; items test word knowledge by selecting 'best-fit' synonym to target words.

3 Administration and Scoring

The MAS range can be administered individually or in a group context. Professionals can choose to administer the MAS online on either the ebilities online platform (with the assistance of the JvR Client Services department) or independently on JvR Online. Scoring is computerised and done automatically after response submission via the ebilities online platform or JvR Online.

3.1 Web-based administration

It is important for assessment administrators to understand how participants complete an online assessment, are able to answer participants' questions or concerns, and can competently use online administrative platforms. The JvR Client Services department can set up individual assessment links on behalf of practitioners wishing to assess candidates on the ebilities online platform, or provide technical assistance to practitioners using the MAS series on JvR Online. JvR Online users will receive detailed instructions on how to use the system and a video tutorial to guide him/her through the process. The JvR Client Services department is also available for additional technical support weekdays from 8am to 5pm (call + 27 11 781 3705 or email clientservices@jvrafrica.co.za).

The JvR Online platform allows users to have full control over their account and offers the following benefits:

- Accessible 24 hours a day, 7 days a week from any PC with internet capability
- Assessment results are available in 90 seconds or less
- Ordering of credits can be done online
- Users can select from the full range of available MAS report options
- Permissions can be set for confidentiality (allowing only particular users access to view and order reports)

4 Interpretation

The MAS range was designed primarily for use in personnel selection, individual assessment, development, and career-related decision-making. The MAS provides information regarding a person's cognitive abilities described by the theories of fluid and crystallised intelligence.

The MAS can be used to assist in the identification of versatile individuals who have the potential to perform at a high level across diverse work tasks. Some of the reports provide a Fundamental Abilities Quotient, which can be used to compare the individual to others from the norm group.

When interpreting the Quotient, scores of 115 and above are considered High. Scores between 86 and 114 are considered Average. Scores of 85 and below are considered Low. Percentages (out of 100) are also reported on for the individual tests.

The MAS reports also provide insight into an individual's self-rated confidence in answering the assessment items correctly, indicating the interaction between ability and confidence in terms of nine Ability Confidence Types:

- **Highly Overconfident** – these individuals display high confidence, but low ability
- **Overconfident** – these individuals display moderate confidence, but low ability
- **Realistic** – these individuals display both low confidence and low ability
- **Underconfident** – these individuals display moderate ability, but low confidence
- **Highly Underconfident** – these individuals display high ability, but low confidence
- **Confident** – these individuals display moderate ability and high confidence
- **Analytical** – these individuals display high ability and moderate confidence
- **Pragmatic** – these individuals display both moderate ability and moderate confidence
- **Incisive** - these individuals display both high ability and high confidence

There are general characteristics, profile strengths and weaknesses associated with High, Moderate, and Low scores. The interpretive statements for each scale are empirically based.

A detailed description of the MAS scales and configurations, along with uses and applications appears in the *ebilities GMA Series Technical Monograph* (Lewis & Cadman, 2017) and *Mental Agility Series Technical Manual* (Douglas, 2017).

5 Feedback and Reporting

Feedback on the MAS range of assessments should be provided on a one-on-one basis. The results should be integrated with those obtained from other assessments and additional information gathered from other sources. Feedback could be in the form of a single session or the results could form the basis of a development plan linked to job objectives. The purpose of the original assessment must be kept in mind when giving feedback. The following result reports are currently available for the MAS assessments:

Table 1 - MAS Battery Compositions

MAS Battery	**Tests Reported On**
General Potential Standard	Numerical Operations Swaps
General Potential Advanced	Series Swaps
Business Fundamentals Standard	Numerical Operations Swaps Vocabulary
Business Fundamentals Advanced	Financial Reasoning Swaps Vocabulary
Numerical Advanced	Financial Reasoning Numerical Operations Series

6 South African Psychometric Properties

The *ebilities GMA Series Technical Monograph* (Lewis & Cadman, 2017) and *Mental Agility Series Technical Manual* (Douglas, 2017) contain reliability and validity information as per the Australian sample. Evidence for the structural, construct, concurrent, and divergent validity of the MAS range of assessments is also available.

The South African sample consisted of 198 working adults who completed the Standard version of the ebilities Swaps, Numerical Operations, and Vocabulary tests between February and November 2018. These scales were chosen to determine how the General Potential Standard and Business Fundamentals Standard in particular may be used within a South African sample. Demographics for this sample can be found in Table 2. Participants' ages ranged between 18 and 65; and 28 individuals did not report their age. Data on language spoken at home was recoded so that Indigenous African Languages formed a group distinct from Afrikaans- and English-speaking individuals. This was to ensure adequate power for subsequent statistical analyses.

Table 2 – Sample Demographics

	N (%)
Gender	
Male	57 (28.8)
Female	141 (71.2)
Highest Qualification	
Below Tertiary	47 (23.7)
Tertiary and Above	151 (76.3)
Employment Position	
Employee/Supervisor	139 (70.2)
Business Owner/Manager/Executive	59 (29.8)
Ethnic Origin	
White	112 (56.6)
Black	35 (17.7)
Coloured	31 (15.7)
Indian/Asian	20 (10.1)
Home Language	
English	108 (54.5)
Afrikaans	58 (29.3)
Indigenous African Languages	31 (15.7)

Note. *N* = sample size; % = percentage of cases in sample.

A statistically significant difference in age was identified between different ethnic groups, $F(3.166) = 4.038$, $p = 0.008$. These findings could largely be accounted for by the observation that individuals reporting their ethnicity as White were older ($M = 37.91$, SD = 9.60) than people who reported their ethnicity as Black ($M = 31.75$, SD = 5.57, $p = 0.001$).

Descriptive statistics for each test can be found in Table 2. The accuracy scores were calculated by summing the result of each item. Reaction times were calculated by averaging the duration in seconds taken to answer each item. Percentage Confidence was calculated by averaging responses to confidence items embedded within each test. All scales had statistics indicating distributions close to normal.

Table 3 - Descriptive Statistics for Swaps, Numerical Operations, and Vocabulary scales

Scale measure	%		N						
	Mean	SD	Mean	SD	α	Min	Max	Skew	Kurt
				Swaps					
Accuracy	65.82	21.70	7.90	2.61	.70	0	100	-.458	-.515
Reaction Time	27.54	6.16			.55	11.79	45.88	.176	.000
Confidence	83.30	14.54			.78	34	100	-.877	.321
				Numerical Operations					
Accuracy	61.73	21.47	9.84	3.43	.74	0	100	-.280	-.305
Reaction Time	14.08	2.10			.33	8.70	20.06	-.106	-.372
Confidence	80.04	17.77			.86	27	100	-.896	.093
				Vocabulary					
Accuracy	75.83	14.21	15.17	2.85	.58	25	100	-.756	.516
Reaction Time	5.82	1.29			.75	3.30	9.55	.455	-.312
Confidence	83.33	13.58			.86	26	100	-1.132	1.389

6.1 Reliability

Internal consistency reliabilities were mostly acceptable for accuracy with the exception of Vocabulary Accuracy. Inspection of the score distributions indicated that this sample scored significantly higher than the Australian norm sample this test was developed in (67.9 percent accuracy versus 75.8 percent accuracy in this sample). These indicative ceiling effects likely reduced the reliability of the scale. All confidence scales demonstrated acceptable internal reliability. As the administration platform is designed to provide cut-off times for responses to questions, the reaction time data likely also suffered from ceiling effects, thereby reducing internal reliability estimates.

6.2 Group Mean Differences

Effect sizes for gender and ethnicity are presented in Tables 4 and 5. We found statistically significant mean differences across both gender and ethnic groups for the tests, making up both the BFQ and GPQ. However, it is not clear at this stage whether these differences are meaningful or simply an artefact of the recruiting strategy for this sample. Further investigation with a larger and more balanced data set would help to clarify this.

6.2.1 Gender

Table 4. Scale Means and Cohen's d Effect Sizes by Gender

	Gender		
Scale	**Men**	**Women**	*d*
Business Fundamentals Quotient	104.96	98.00	0.48*
General Potential Quotient	105.40	97.82	0.52*
Swaps	67.33	65.21	0.10
Numerical Operations	70.16	58.33	0.57*
Vocabulary	80.00	74.15	0.43*

Note. Cohen's d Effect size 0.20-0.49 = small, 0.50-0.79 = medium, 0.80 and above= large. Independent sample t tests were used to determine group differences.
**p < 0.05.*

Independent samples *t* tests revealed statistically significant gender differences between women and men, with women scoring lower on both the Business Fundamentals Quotient and the General Potential Quotient. Running the same analyses on the constituent tests revealed no statistically significant differences between women and men on Swaps, indicating that there are negligible differences in working memory performance between men and women. However, statistically significant differences between women and men were identified on both Numerical Operations and Vocabulary, with women scoring significantly lower than men on both tests. When considering effect sizes, the most significant differences in performance was observed on the Numerical Operations test, which was part of both the BFQ and GPQ assessments.

However, these results should be read with caution given the limited male sample that represented under 30 per cent of the overall sample that was used for this analysis.

6.2.2 Ethnicity

Table 5. ANOVA η_p^2 Effect Sizes by Ethnicity

	Group Comparisons				
Scale	***df***	***F***	***p***	η_p^2	**Direction**
Business Fundamentals Quotient	3.194	12.940	< .001	.167	1,3,4 > 2
General Potential Quotient	3.194	8.050	< .001	.111	
Swaps	3.194	12.762	< .001	.064	
Numerical Operations	3.194	4.423	.005	.005	
Vocabulary	3.194	11.169	< .001	.147	

Note. 1= White; 2= Black; 3= Coloured; 4 = Indian/Asian. Effect size: 0.01-0.05 = small, 0.06-0.13 = medium, 0.14 and above = large.

Individuals who reported Black ethnic origins, scored lower on both the BFQ and GPQ than all other ethnic groups. Results were further examined at the individual test level, and findings indicated lower scores for Blacks on all three tests, who scored statistically significantly lower than both White and Indian/Asian individuals, with the difference between Black and Coloured participants trending towards significance.

In relation to Ethnicity, we found that the largest difference between groups was due to performance on the Vocabulary test. This is to be expected given that just over half of the sample were English home language speakers, and that within the South African context, many speak English as a second, or even third language. In contrast to the gender comparison, there was a significant difference in how different ethnicities performed on the Swaps test, with the smallest difference reported on the Numerical Operations test.

Again, it must be noted that Black participants were under-represented and made up under 20 percent of the overall sample, and therefore these analyses should be interpreted with caution.

7 Construct Validity/Correlations between Scales

The theory of intelligence suggests that higher accuracy scores should correspond with faster processing speeds. To confirm the theoretical underpinning of the batteries, principal component analyses were conducted to investigate whether the accuracy scores and reaction time scores for the various scales could be appropriately combined into a single score to determine either a Business Fundamentals Quotient (BFQ) or General Potential Quotient (GPQ).

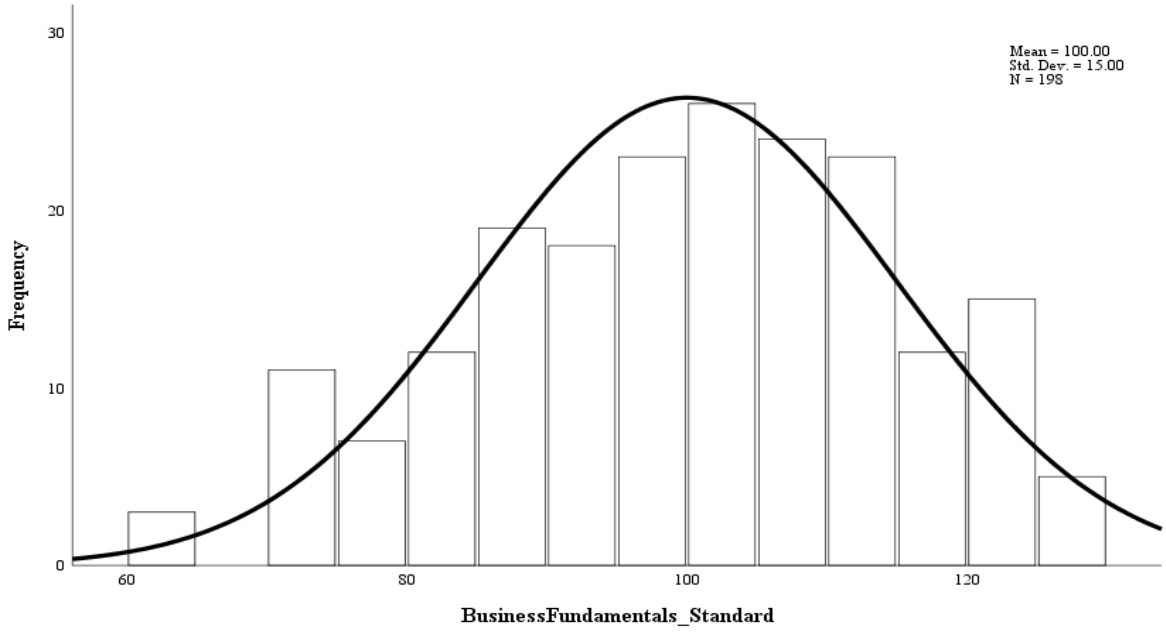
7.1 Business Fundamentals Standard

In the Business Fundamentals Standard report, accuracy scores for each of the three scales, as well as the recorded reaction times, are used to calculate a Business Fundamentals Quotient (BFQ). Confidence scores across the three scales determine the overall Confidence range reported on.

7.1.1 Business Fundamentals Quotient (BFQ)

A principal components analysis indicated that the accuracy scores for Swaps, Numerical Operations, and Vocabulary, and the reaction time scores for Numerical Operations and Vocabulary, could be appropriately combined into a single score. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analyses, KMO = 0.610. Bartlett's test of sphericity $\chi 2(10) = 226.726$, $p < 0.001$, confirmed that the correlations between items were large enough for factor analyses. The preferred one-component solution accounted for 44.86 per cent of the variance in the data. Communalities for the one-component solution ranged between 0.275 and 0.610, with all five variables loading on the single component above 0.50. All five variables were therefore included in the calculation of the BFQ, and the distribution of the standardised BFQ can be found in Figure 1.

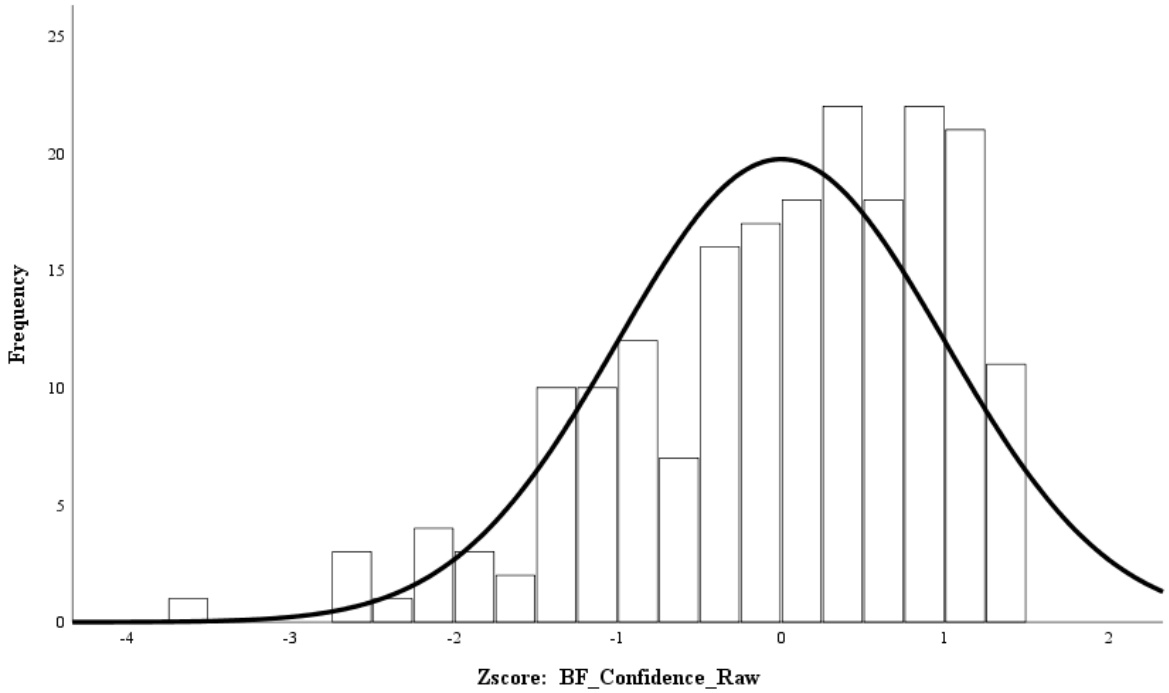
Figure 1. Distribution of BFQ



7.1.2 Business Fundamentals Confidence Score

The confidence scores from Swaps, Numerical Operations, and Vocabulary were submitted to a principal components analysis to determine the appropriate scoring of Confidence associated with the Business Fundamentals. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analyses, KMO = 0.631, and Bartlett's test of sphericity $\chi2(3) = 118.657$, $p < 0.001$, confirmed that the correlations between items were large enough for factor analyses. The preferred one-component solution accounted for 62.94 per cent of the variance in the data. Communalities for the one-component solution ranged between 0.481 and 0.723, with all three variables loading on the single component above 0.69. The distribution of the Business Fundamentals Confidence score can be found in Figure 2.

Figure 2. Distribution of Business Fundamentals Confidence Score



7.2 General Potential Standard

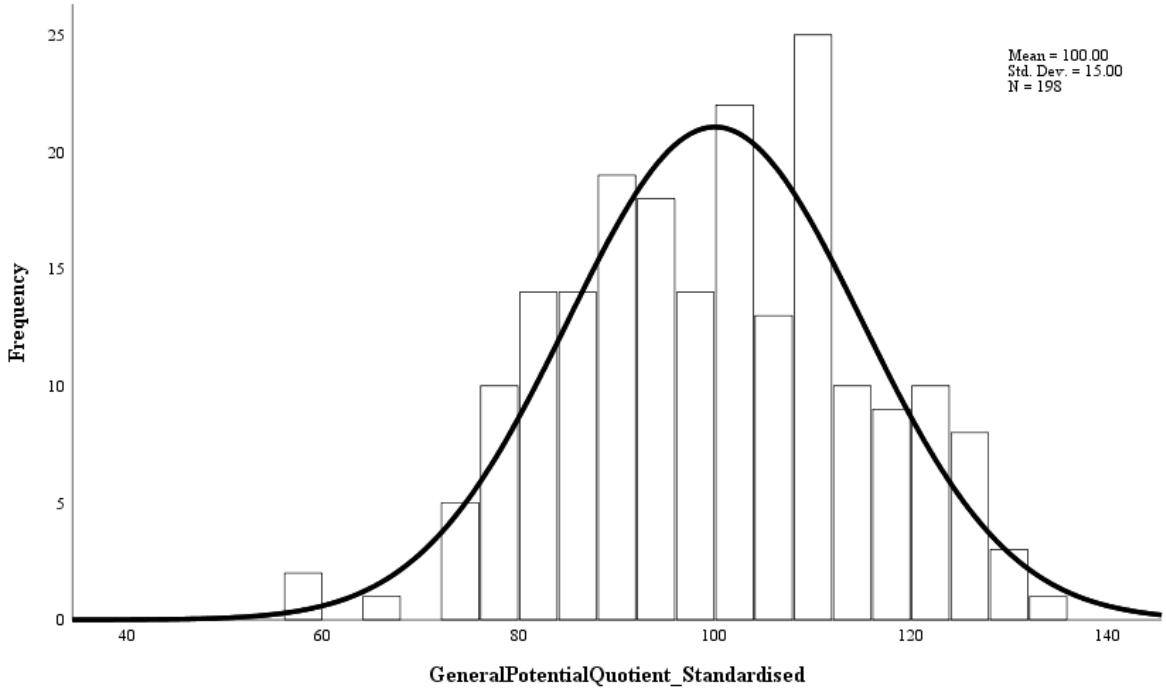
In the General Potential Standard report, accuracy scores for each of the two scales, as well as the recorded reaction times, are used to calculate a General Potential Quotient (GPQ). Confidence scores across both scales determine the overall Confidence range reported on.

7.2.1 General Potential Quotient (GPQ)

Accuracy data for Numerical Operations and Swaps, and reaction time data for Numerical Operations, were entered into a principal components' analysis, which confirmed that the correlations between items were large enough for factor analyses. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analyses, KMO = 0.548, Bartlett's test of sphericity $\chi2(3) = 68.329$, $p < 0.001$. The preferred one-component solution accounted for 54.094 per cent of the variance in the data. Communalities for the one-component solution ranged between 0.262 and

0.711, with all three variables loading on the single component above 0.50. The distribution of the standardised GPQ can be found in Figure 3.

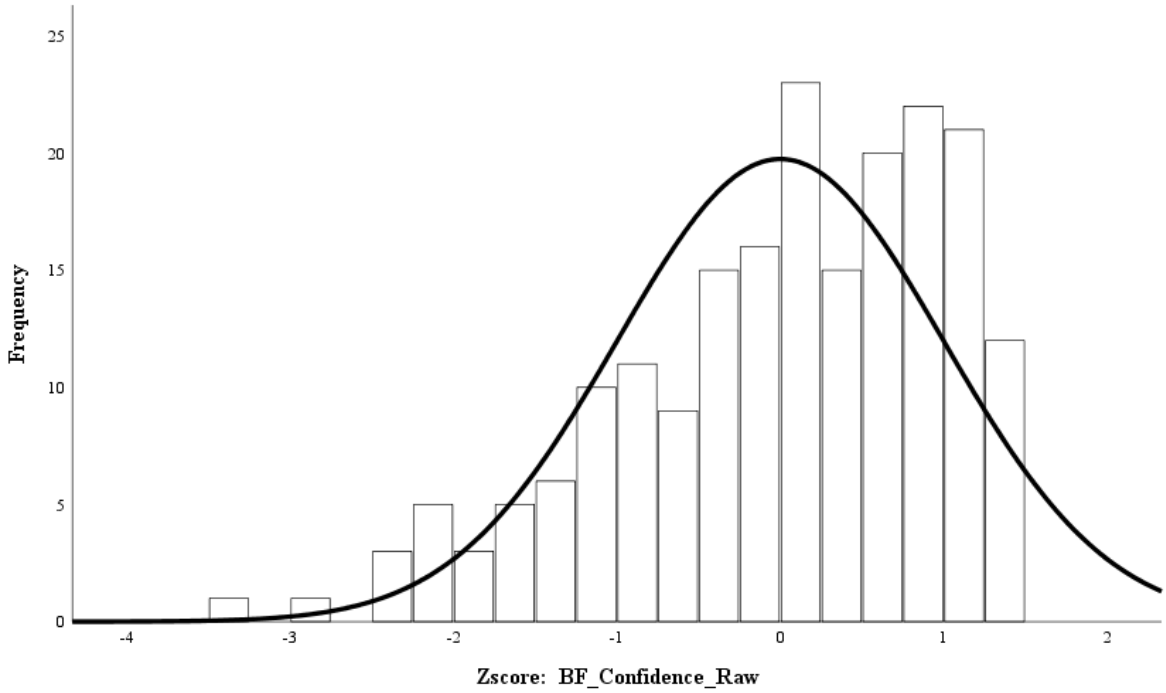
Figure 3. Distribution of GPQ



7.2.2 General Potential Confidence Score

The Confidence scores for Swaps and Numerical Operations were subjected to a principal components analysis to determine the appropriate scoring of Confidence associated with General Potential scores of individuals. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analyses, KMO = 0.500. Bartlett's test of sphericity $\chi 2(1) = 81.727$, $p < 0.001$, confirmed that the correlations between items were large enough for factor analyses. The preferred one-component solution accounted for 79.23 per cent of the variance in the data. Communalities for the one-component solution were both 0.792, with both variables loading on the single component above 0.80. The distribution of the General Potential Confidence score can be found in Figure 4.

Figure 4. Distribution of General Potential Confidence Score



9 Concluding Comments

South African norms for the Business Fundamentals Standard and General Potential Standard, as well as the tests comprising them, were generated based on the data collected from 198 working South African adults during the period February to December 2018. Overall, the psychometric properties were acceptable, and the assessment appears to be appropriate for use in South African samples. Although there appear to be statistically significant gender and ethnic differences, it is recommended that these be analysed with caution, given the small sample size and noting that White females with tertiary education were over-represented in the sample. The largest differences in the sample also appeared to be on their Vocabulary performance, where it is expected that English first-language speakers would perform better. As more data is collected on the assessment, different norms will be created for English first-language and English second-language speakers over a General South African norm.

10 References

Douglas, H. (2017). *Mental agility series technical manual: Test properties, short forms and test combinations.* Surry Hills, Australia: Lewis Cadman Consulting Pty Ltd.

Lewis, C., & Cadman, M. (2017). *ebilities GMA series technical monograph*. Surry Hills, Australia: Lewis Cadman Consulting Pty Ltd.