SOUTH AFRICAN RESEARCH SUPPLEMENT

AUGUST 2018

Authors:

C. Hayes

Dr N. Taylor

S. Bakker

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Career Values Scale (CVS) was developed by Psychometrics Canada and is a measure of work values, preferences, and needs. Because the work values form a part of an individual's core beliefs and give meaning to their career, they are seen as a useful indicator of job satisfaction and personal goals.

The CVS consists of 10 scales (see Table 1) and 88 items, with each one of the scales having between 7 and 12 items. Participants are asked to rate the items on a five-point Likert type scale ranging from not at all important, to very important:

1. Not at all important
2. Unimportant
3. Neutral
4. Important
5. Very important.

The Importance of Career Values

Values as a construct has been researched by many scholars (Schwartz, 1999; Gouveia, Millfont, & Guerra, 2014) internationally, and within the South African context (Coetzee, Schreuder& Tladinyane, 2014; Du Toit & Coetzee, 2012; Neelam & Tanksale, 2014). Values can be seen as behavioural moderators that are formed as an expression of personal needs, beliefs, and attitudes (Schwartz, 2002, Gouveia, et al., 2014, Neelam & Tanksale, 2014), and can occur on a personal or collective level, including societal norms, organisational values, or professional standards.

The importance of values within the career context has been well documented, with scholars arguing that a fit between an individual's values and that of an organisation, can lead to higher job satisfaction (Diskiené & Goštatus, 2013), greater productivity/performance (Dearlove and Coomber, 1999), and higher commitment from employees (Sullivan, Sullivan and Buffton, 2002), to mention a few. Literature further suggests that career choices made by individuals are influenced by their personal values (Schreuder & Coetzee, 2011; Schwartz, 1999). Miller & Miller (2005) further also suggest that understanding how an individual's personal values relate to their career choices form an important underlying factor in career development.

Being able to measure to what extent individuals value specific domains, can give career counsellors insight into which careers or working environments would be best suited for the individual. It can further aid individuals in their personal career development by gaining an understanding of which areas or levels of work are aligned to their own career values.

Scales

In Table 1 we provide a breakdown of the ten career values, with a description of what a high score in each scale indicates.

Table 1. Scale name and definitions.

Factor	Scale (value)	High end scale descriptions*
Working with others	Service Orientation	People with high scores on *Service Orientation* value helping people with their problems. They place a high value on work settings where you can gauge the thoughts and feelings of people. Common examples include care giving and customer service positions. They value activities where there is a direct positive effect upon others.
	Team Orientation	People with a high score on *Teamwork* value working closely with co-workers and clients. They value encouragement and encouraging others. Occupations that rely heavily on teamwork, networking, and relationships will be rewarding.
	Influence	People with a high score on *Influence* value being directive rather than collaborative, and like to control and manage the work that they are involved with. They will enjoy having a high level of responsibility, and will be satisfied in an organisation where they can use a directive leadership style.
Self-expression	Creativity	People with high scores on *Creativity* value working in an organisation that promotes ingenuity and creativity. They value developing new ideas, exploring unconventional approaches, and using their imagination. The types of work that will be satisfying will require creativity and innovation, and will allow the client to use their curiosity to identify and evaluate new ideas.
	Independence	People with high scores on *Independence* value the self-reliance and the freedom to set their own goals and schedule. They value working without supervision, and enjoy tasks that can be worked on independently without having to seek advice from co-workers. The type of work they enjoy will allow them to do things in their own way, most of the time.
	Excitement	People with high scores on *Excitement* value trying new things. They value environments where there is uncertainty and where change is expected. The type of work they will enjoy will include the possibility of trying something new, and where they can take risks.
	Career Development	People with high scores on *Career Development* value developing both personally and professionally. They value learning new skills and expanding their expertise, and will enjoy working in an organisation where there is a culture of life-long learning, where there are constant challenges, and where they can have an opportunity to develop new skills.
Extrinsic rewards	Financial Rewards	People with high scores on *Financial Rewards* value having enough income to ensure that they can afford life's luxuries. Excellent financial rewards are the key to their happiness and satisfaction. They value working in an organisation that values financial incentives, and where they are paid according to their performance.
	Prestige	People with high scores on *Prestige* value recognition and acclaim. They value working in an organisation where they are publicly recognised as a superior performer. They value working for an organisation that in itself is well-known or is recognised as being amongst the best.
	Security	People with high scores on *Security* value a consistent career path and a feeling of job security. They prefer the steady, and the predictable, to rapid and unpredictable change, and will enjoy working where they feel that they can plan for the future.

Note: * - taken from McNab, Bakker & Fitzsimmons (2005)

Administration and Use

In the South African context, the CVS can only be used by registered psychology professionals.

Age

The CVS can be used with participants who are 15 years and older. It is however recommended that it only be used with those individuals with work experience.

Reading level

Participants require a reading ability compatible to Grade 8 (NQF Level 1).

Areas of application

The CVS can be used in many career-related spheres, including career counselling, team building, and organisational planning.

Web administration

The CVS is only available online, and clients can access it through the JvR Online portal.

Training

No accreditation training is required for the CVS.

Development

For in-depth information on the development of the CVS, please refer to the Career Values Scale Manual and User Guide (McNab, et al., 2005).

SAMPLE DESCRIPTION

This section provides an overview of the current sample used for the South African research supplement. Table 2 provides information on the various age, gender, and ethnic groups, included in the sample. The sample was obtained from the online portal, and consisted of individuals who have completed the assessment between 2011 and 2017. All participants completed online consent for their data to be used for further research on the instrument.

Table 2. Sample distribution.

Group		*N*	%
Age	15-20	351	16.3%
	21-30	162	7.5%
	31-40	108	5.0%
	41-50	69	3.2%
	51-60	25	1.2%
	Missing	1439	66.8%
		2154	**100.0%**
Gender	Men	1043	48.4%
	Women	1111	51.6%
	Missing	0	0.0%
		2154	**100.0%**
Ethnicity	Asian/Indian	99	4.6%
	Black	426	19.8%
	Coloured	42	1.9%
	White	147	6.8%
	Other	4	.2%
	Missing	1436	66.7%
		2154	**100.0%**

The sample consisted of a relatively equal representation of men and women. The largest portion of participants (>66%) did not indicate their ethnicity or age, but of the participants who did, the largest sample group was black participants, and participants between the ages of 15-20 of which the majority were between 18 -20 years (*N*=335).

DESCRIPTIVE STATISTICS

In this section we provide an overview of the overall sample's performance on the CVS, with tests of normality for each subscale.

Overall performance

Table 3 provides an indication of the overall sample performance on each of the ten CVS scales. From the table it can be determined that the current sample indicates that the overall sample performed within the average range on all of the individual CVS scales. When converting the sample's average scores for each scale, the overall sample obtained a Sten of 5 for each of the CVS scales (based on the US norms).

Table 3. Overall sample scale scores.

Scale	*N*	Mean	S.D.	S.E.	Min	Max	Range
Service Orientation	2154	29.28	3.601	0.078	13	35	22
Team Orientation	2154	43.65	5.642	0.122	21	55	34
Influence	2154	35.19	5.099	0.110	16	45	29
Creativity	2154	31.96	5.044	0.109	11	40	29
Independence	2154	36.43	5.062	0.109	19	50	31
Excitement	2154	24.62	3.799	0.082	11	35	24
Career Development	2154	38.50	4.199	0.090	19	45	26
Financial Rewards	2154	42.28	6.237	0.134	18	55	37
Prestige	2154	29.55	5.037	0.109	10	40	30
Security	2154	26.33	3.584	0.077	8	35	27

Overall sample distribution

The scores distribution for each of the CVS scales were negatively skewed, indicating that participants from the current sample tended to score toward the higher end of the score range. In order to investigate if the skewness was statistically significant, we ran a Shapiro-Wilk test which yielded significant results for all the scales. With larger sample sizes there is a risk that even minimal deviations from the normal distribution will be flagged as significant. In order to verify the findings, we further investigated the skewness statistics, potential outliers (box plots), and the Q-Q plots for each scale. Service Orientation, Creativity and Career Development were all moderately skewed (<-0.50), while the rest of the scales' skewness statistics indicated approximately symmetrical distributions. When investigating the box plots for each scale, several outliers where identified, but these were all found to be valid representations of individual scores. The Normal Q-Q Plots for the various scales indicated only a slight deviance from the normal distribution.

Although none of the scales met the requirements for normal distribution based on the tests of normality, due to the large sample sizes, and the inspection of the Q-Q Plots, that only showed slight deviance, we conducted both non-parametric and parametric analyses, with the latter being reported in the document.

Table 4. Sample distribution descriptive statistics

Scale	Skewness	S.E.	Kurtosis	S.E.	Test for Normality*		
					Statistic	df	Sig.
Service Orientation	-.710	.053	.732	.105	0.960	2154	.000**
Team Orientation	-.410	.053	.362	.105	0.984	2154	.000**
Influence	-.361	.053	-.102	.105	0.985	2154	.000**
Creativity	-.622	.053	.307	.105	0.967	2154	.000**
Independence	-.116	.053	-.042	.105	0.995	2154	.000**
Excitement	-.038	.053	.242	.105	0.992	2154	.000**
Career Development	-.736	.053	.711	.105	0.957	2154	.000**
Financial Rewards	-.377	.053	.029	.105	0.987	2154	.000**
Prestige	-.263	.053	-.125	.105	0.990	2154	.000**
Security	-.487	.053	.581	.105	0.981	2154	.000**

Note: * - Shapiro-Wilk test for normality; ** - statistically significant at p<0.05

RESPONSE STYLE

We decided to explore the response style of the current sample in order to investigate whether respondents were more inclined to respond toward the extreme ends of the scale (i.e., selecting 1 or 5). Figure 1 provides an overview of how the total sample responded to the items in the CVS. As is evident from the results, most participants in the current sample tended to endorse the higher spectrum responses (4-5; important to very important). This is congruent with the skewness results. This pattern will be further investigated in the section addressing results from the Rasch analysis.

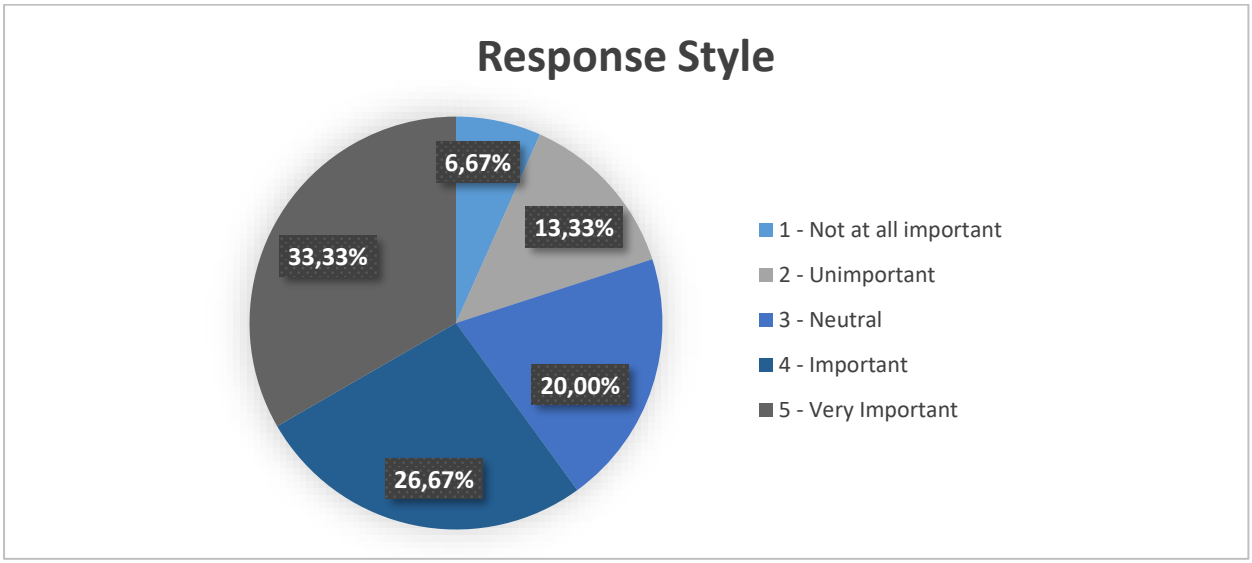

Figure 1. Response style

RELIABILITY

Overall sample

In Table 5, two measures of internal consistency reliability are reported, namely Cronbach's Alpha and Guttman's Lambda 2. Except for the Security and Excitement scales, all the other CVS scales yielded good internal consistency reliability (> 0.70). It is noteworthy that these two scales only consist of 7 items, which could influence their reliability. Some scholars argue that reliability estimates above 0.60 are still acceptable in self-report assessments, although the results should not be used on their own to inform final decisions (Urbina, 2004). An overall reliability of the CVS is not reported as there is no overall score on the CVS, but only scale scores that are interpreted.

Table 5. Reliability for the overall sample

Scale	Items	α	λ2
Service Orientation	7	.805	.809
Team Orientation	12	.825	.830
Influence	9	.833	.835
Creativity	8	.864	.867
Independence	10	.704	.714
Excitement	7	.683	.697
Career Development	9	.790	.798
Financial Rewards	11	.825	.832
Prestige	8	.780	.789
Security	7	.667	.675

We inspected the inter-item correlations and found moderate to high correlations between items (0.10-.060). Some of the scales yielded smaller correlations between items (< 0.08), and in some cases there were items in specific scales that correlated only slightly with all other items in the scale. One item from the Independence scale also correlated negatively with other items in that scale. These items are flagged for further analyses in the factor and Rasch analyses phases, as they may be measuring different constructs to what they purport to measure. A list of these items is given in Table 6. We also inspected the item total statistics to identify any items that might potentially be inflating or reducing the internal consistency of the scales. Item CAD1 from the Career Development scale was identified as a potentially problematic item, as the removal of this item would increase the reliability of the scale with 0.079. No further items were identified based on the item total statistics.

Table 6. Items identified for further investigation.[1]

Scale	Item number
Independence	IND1 and IND6
Career Development	CAD1

[1] Abbreviations of the scale names appear with each item in order to help identify the item: Service Orientation – SO, Team Orientation – TO, Influence – INF, Creativity – CRE, Independence – IND, Excitement – EXC, Career Development – CAD, Financial Rewards – FIR, Prestige – PR, and Security – SEC.

Sample subgroups

In order to investigate the reliability of the CVS more thoroughly, we calculated internal consistency reliability coefficients for specific subgroups within the sample. We looked at age, gender, and ethnic groupings.

Age

Table 7 gives an overview of the Cronbach's alpha and Guttman's Lambda 2 estimates of internal consistency reliability for the different CVS scales across the six defined age groups.

Table 7. Reliability for age groups

		15-20		21-30		31-40		41-50		51-60	
Scale	**Items**	**α**	**λ2**	**α**	**λ2**	**α**	**λ2**	**α**	**λ2**	**α**	**λ2**
Service Orientation	7	.740	.746	.768	.746	.786	.795	.807	.815	.823	.852
Team Orientation	12	.789	.796	.808	.816	.851	.861	.815	.825	.871	.885
Influence	9	.708	.714	.756	.763	.849	.855	.773	.784	.822	.844
Creativity	8	.742	.749	.836	.844	.898	.902	.844	.853	.883	.894
Independence	10	.590	.601	.702	.716	.719	.734	.773	.790	.683	.731
Excitement	7	.552	.579	.573	.596	.740	.755	.711	.734	.696	.727
Career Development	9	.594	.612	.731	.748	.709	.718	.749	.756	.664	.728
Financial Rewards	11	.779	.787	.809	.822	.819	.829	.803	.817	.819	.842
Prestige	8	.691	.702	.716	.724	.713	.740	.851	.873	.804	.831
Security	7	.466	.480	.509	.550	.786	.792	.788	.793	.806	.815

The scales from the "Working with others" domain, along with Creativity and Financial Rewards, showed good internal consistency reliability (>0.70) across all of the age groups, but showed lower reliability with the 15-20-year-old age group than any of the other age groups.

Independence, Career Development, and Prestige, showed lower internal consistency (<0.70) for the 15-20-year-old age groups, but good internal consistency reliability for the rest of the age groups. The Excitement and Security scales yielded lower reliability for the two younger age groups (<0.60), but showed good internal consistency reliability for age groups above 31-years of age.

These results suggest that some of the CVS scales are not appropriate for use with younger population groups who do not have the necessary exposure to working environments.

Gender

Seven of the CVS scales showed good internal consistency (>0.70) for both men and women. The Excitement and Security scales had lower reliability coefficients (<0.70) for both gender groups, and the Independence scale had good internal consistency for men, but lower internal consistency for women, in the current sample. All of the scales still yielded internal consistency ranging between 0.64 and 0.88, indicating acceptable to good reliability for the CVS for both gender groups. Table 8 gives an overview of both the Cronbach's alpha and Guttman's Lambda 2 measures of internal consistency across the 10 CVS scales for both gender groups.

Table 8. Reliability for gender groups

		Men		Women	
Scale	Items	α	λ2	α	λ2
Service Orientation	7	.810	.813	.802	.806
Team Orientation	12	.837	.842	.814	.820
Influence	9	.836	.839	.825	.827
Creativity	8	.848	.850	.871	.875
Independence	10	.713	.724	.694	.704
Excitement	7	.699	.711	.673	.688
Career Development	9	.789	.796	.791	.800
Financial Rewards	11	.820	.828	.831	.838
Prestige	8	.783	.792	.776	.786
Security	7	.684	.692	.645	.656

Ethnicity

The results from the reliability analyses for the different ethnic groups are reported in Table 9. Cronbach's alpha and Guttman's Lambda 2 are reported.

Table 9. Reliability for ethnic groups

		Asian/Indian		Black		Coloured		White	
Scale	Items	α	λ2	α	λ2	α	λ2	α	λ2
Service Orientation	7	.760	.776	.756	.762	.763	.773	.779	.785
Team Orientation	12	.800	.811	.773	.779	.810	.831	.834	.847
Influence	9	.789	.799	.748	.753	.785	.799	.846	.852
Creativity	8	.796	.804	.795	.801	.744	.766	.801	.809
Independence	10	.646	.679	.646	.657	.634	.662	.709	.723
Excitement	7	.451	.502	.579	.607	.582	.611	.743	.757
Career Development	9	.739	.747	.631	.645	.672	.711	.759	.766
Financial Rewards	11	.673	.694	.798	.808	.784	.798	.839	.846
Prestige	8	.771	.791	.675	.690	.765	.790	.761	.781
Security	7	.726	.737	.583	.593	.644	.668	.735	.742

The scales related to "Working with others" and the Creativity scale, showed good reliability across all of the ethnic groups in the current sample. All of the CVS scales had good internal consistency for the White population group (>0.70), with all but two scales (Independence and Excitement) yielding reliability coefficients above 0.70 for the Asian/Indian population group.

Three scales from the "Self-expression" domain (Independence, Excitement and Career Development) had reliability coefficients below 0.70 for the Asian/Indian, Black, and Coloured population groups, with Excitement yielding internal consistency values below 0.60 for both the Asian/Indian and Black population groups.

The results are in line with the overall reliability scale results, showing slightly lower internal consistency reliability with Excitement and Security scales.

GROUP COMPARISONS

In order to investigate how the CVS scales perform within different samples we ran group comparisons (*t*-tests and ANOVAs) on three specific sample subgroups: age, gender, and ethnicity. The results from these analyses are presented below.

Age

We provide an overview of the performance of different age groupings on the CVS in terms of their mean scores and standard deviations in Table 10. It is important to note that certain age groups had significantly more participants than other groups.

Table 10. Mean scores for age groups

	15-20 (*N* = 351)		21-30 (*N* = 162)		31-40 (*N* = 108)		41-50 (*N* = 69)		51-60 (*N* = 25)	
Scale	***M***	**SD**	***M***	**SD**	***M***	**SD**	***M***	**SD**	***M***	**SD**
Service Orientation	29.34	3.70	28.75	3.64	29.98	2.91	30.62	2.83	31.12	2.80
Team Orientation	44.23	5.67	43.33	5.88	42.54	5.47	44.16	4.53	44.16	5.58
Influence	34.20	4.71	34.74	4.69	37.21	4.42	38.55	3.59	39.28	3.79
Creativity	33.57	4.18	32.56	4.77	32.63	4.84	32.45	4.34	34.44	4.11
Independence	38.42	4.70	37.46	5.27	35.42	4.95	35.70	5.28	35.20	4.58
Excitement	24.66	3.78	25.21	3.50	24.24	3.77	25.19	3.35	25.68	3.49
Career Development	38.16	3.60	38.14	4.22	39.55	3.22	39.32	3.36	39.60	3.20
Financial Rewards	44.15	6.20	42.94	6.28	42.05	5.53	41.86	5.40	39.00	5.83
Prestige	30.44	4.85	29.70	4.81	28.29	4.23	28.67	5.67	29.84	5.10
Security	27.82	3.11	26.58	3.38	25.70	4.02	24.20	4.30	24.00	4.37

Table 11 provides the results from the one-away analysis of variance conducted for the different age groups. Results indicated that there were statistically significant differences between age group performances on eight of the CVS scales. The only scales were there were no statistically significant differences between age group scores, were the Team Orientation and the Excitement scales.

Table 11. ANOVA output for age groups

Scale	***df***	***F***	***p***	n_p^2
Service Orientation	4, 710	5.769	.000	.180
Team Orientation	4, 710	2.264	.061	.113
Influence	4, 710	24.244	.000	.369
Creativity	4, 710	2.831	.024	.126
Independence	4, 710	11.513	.000	.255
Excitement	4, 710	1.822	.123	.101
Career Development	4, 710	4.764	.001	.164
Financial Rewards	4, 710	7.089	.000	.200
Prestige	4, 710	5.101	.000	.170
Security	4, 710	23.785	.000	.366

In order to determine where the significant differences existed for the various age groups, we ran post-hoc tests. These results are presented in Table 12. Note that only the statistically significant differences are reported in the table. Tukey's honest significant difference test was run and reported on samples that had equal variances; and Games-Howell was run and reported for the scales that did not meet the requirements of homogeneity of variance. The results of the post-hoc tests are discussed for each scale.

Table 12. Post-hoc tests results for age groups

								95% CI	
Scale	Post-hoc Test	Age Groups		Mean Difference	S.E.	Sig.	*d*	Lower Bound	Upper Bound
Service Orientation	Games-Howell	15-20	41-50	-1.287	.458	.040	.39	-2.54	-.04
		21-30	31-40	-1.235	.432	.035	.37	-2.42	-.05
			41-50	-1.876	.500	.002	.57	-3.24	-.51
			51-60	-2.373	.747	.013	.73	-4.42	-.33
Team Orientation	Tukey HSD	15-20	31-40	1.688	.615	.049	.30	.01	3.37
Influence	Tukey HSD	15-20	31-40	-3.016	.499	.000	.66	-4.38	-1.65
			41-50	-4.354	.597	.000	.66	-5.99	-2.72
			51-60	-5.083	.939	.000	1.04	-7.65	-2.52
		21-30	31-40	-2.472	.563	.000	.54	-4.01	-.93
			41-50	-3.810	.652	.000	.91	-5.59	-2.03
			51-60	-4.539	.974	.000	1.06	-7.20	-1.87
Independence	Tukey HSD	15-20	31-40	3.008	.541	.000	.62	1.53	4.49
			41-50	2.729	.648	.000	.54	.96	4.50
			51-60	3.225	1.019	.014	.69	.44	6.01
		21-30	31-40	2.046	.611	.008	.40	.37	3.72
Career Development	Tukey HSD	15-20	31-40	-1.387	.403	.006	.41	-2.49	-.28
		21-30	31-40	-1.404	.455	.018	.38	-2.65	-.16
Financial Rewards	Tukey HSD	15-20	31-40	2.099	.664	.014	.36	.28	3.92
			41-50	2.290	.795	.033	.39	.12	4.46
			51-60	5.145	1.249	.000	.86	1.73	8.56
		21-30	51-60	3.938	1.297	.021	.65	.39	7.48
Prestige	Tukey HSD	15-20	31-40	2.152	.533	.001	.47	.69	3.61
			41-50	1.772	.638	.045	.34	.03	3.52
Security	Games-Howell	15-20	21-30	1.237	.313	.001	.38	.38	2.10
			31-40	2.114	.421	.000	.59	.95	3.28
			41-50	3.615	.543	.000	.97	2.10	5.13
			51-60	3.818	.889	.002	1.01	1.21	6.42
		21-30	41-50	2.377	.581	.001	.62	.76	3.99

Service Orientation

The 21–30-year-old group scored statistically significantly lower on Service Orientation than all of the groups older than them. This would suggest that younger participants starting out in their careers place less value in service to others, than participants with more working experience. The 15–20-year-old age group also scored statistically significantly lower than the 41–50-year-old participants, further showing that older participants with more work experience tend to value Service Orientation more than their younger counterparts. The effect sizes for the differences were small to moderate (0.37-0.73), suggesting that these differences are meaningful and age specific norms should be considered.

Team Orientation

15–20-year-old participants scored statistically significantly higher than 31–40-year-olds on the Team Orientation scale (*d*=0.30), suggesting that younger participants with less working experience might value teamwork more than those in their middle career stages. No other statistically significant differences exist between different age groups on this scale.

Influence

The two younger age groups scored similarly on the Influence scale, and both scored statistically significantly lower than all three of the older age groups, with moderate to large effect sizes (0.54–1.06). This would suggest that Influence is valued more by participants who have progressed more in their careers, or have more experience within the career context. These differences suggest that it might be necessary to consider age specific norms for the CVS.

Creativity

Although the ANOVA revealed that there were statistically significant differences between age groups on the Creativity scale, the post-hoc tests did not reveal any statistically significant differences.

Independence

On the Independence scale the post-hoc results indicated that the 15–20-year-old group scored statistically higher than the 31–40, 41–50, and 51–60-year-old groups (*d*:0.54-0.69), while the 21–year-old participants also scored statistically significantly higher than 31–40-year-old participants (*d*=0.40). The moderate effect sizes of these differences suggest that age-specific norms may be more appropriate for the CVS than a general norm sample.

Career Development

On the Career Development scale, the 31–40-year-old group scored significantly higher than the 15–20 and 21–30-year-old groups, with moderate effect sizes. This would suggest that participants entering the middle stages of their careers place more value in developing their careers, than those with minimal working experience or just starting out in their respective careers.

Financial Rewards

Younger participants (15–20-year-old) scored statistically significantly higher on Financial Rewards than participants from the 31–60-year-old groups. 51–60-year-old participants also scored statistically significantly lower than 21–30-year-old participants on this scale. This would suggest that Financial Rewards are valued more by younger participants with less working experience, and that the oldest age group in the sample places less value in this than their younger counterparts.

Prestige

The 15–20-year-old participants scored statistically significantly higher on Prestige than participants from the 31–40 and 41–50-year-old groups, yielding moderate effect sizes. This suggests that

younger participants who do not have a lot of working experience might value recognition and acclaim more than those participants who are already settled in their careers.

Security

Post-hoc results revealed that participants from the 15–20-year-old group scored significantly higher than all of the other age groups on the Security scale. The moderate to large effect sized in these differences are expected within the South African context, with high unemployment rates that could impact on the youth and their career choices. Early career participants (21–30-year-olds) also scored statistically significantly higher than 41–50-year-old participants.

Summary

The results from the ANOVA run on age group differences suggest that it might be necessary to have age-specific norm groups for the interpretation of CVS scores. It could, however, also highlight that the younger age group's (15-20) results could be influenced by the lack of work experience, or exposure to the job market that participants under the age of 18 might have. This group had the most statistically significant differences from other age groups within the current sample. It could be indicative that the CVS is not applicable to population groups younger than 18-years of age.

Gender

We provide an overview of the mean scores for each gender group on the different CVS scales in Table 13. The results from independent t-tests that were run on the two gender groups are presented in Table 14.

Table 13. Mean scores for gender groups

	Men			Women		
Scale	***N***	***M***	**SD**	***N***	***M***	**SD**
Service Orientation	1043	29.12	3.503	1111	29.43	3.686
Team Orientation	1043	43.66	5.699	1111	43.65	5.591
Influence	1043	35.95	4.834	1111	34.48	5.239
Creativity	1043	32.58	4.542	1111	31.38	5.411
Independence	1043	35.98	5.115	1111	36.85	4.978
Excitement	1043	24.75	3.754	1111	24.51	3.839
Career Development	1043	38.68	4.043	1111	38.33	4.336
Financial Rewards	1043	42.53	5.999	1111	42.05	6.446
Prestige	1043	29.74	4.934	1111	29.38	5.128
Security	1043	26.02	3.650	1111	26.62	3.496

Table 14. T-test results for gender comparisons

	Levene's Test		t-test for Equality of Means			
Scale	***F***	**Sig.**	***t***	**df**	**Sig. (2-tailed)**	***d***
Service Orientation	2.684	.102	-2.018	2152	.044*	.09
Team Orientation	.037	.848	.043	2152	.966	.00

Influence	10.117	.001*	6.744	2151.369	.000*	.29
Creativity	30.940	.000*	5.627	2125.794	.000*	.24
Independence	.160	.689	-3.981	2152	.000*	.17
Excitement	.767	.381	1.495	2152	.135	.06
Career Development	1.677	.195	1.930	2152	.054	.08
Financial Rewards	5.998	.014*	1.807	2151.841	.071	.08
Prestige	.818	.366	1.690	2152	.091	.07
Security	.728	.394	-3.896	2152	.000*	.23

There were statistically significant differences across genders for five of the CVS scales. Women scored statistically significantly higher than men on Service Orientation, Independence, and Security, but the effects of these differences were small (0.09-0.23). Men scored significantly higher than women on Influence and Creativity, with small effect sizes (0.17-0.29).

Summary

Although differences exist between gender groups on five of the CVS scales, the effect sizes of these differences are seen as small, and it is argued that the sample size could have caused these small differences to be inflated to statistical significance. At this stage there is insufficient evidence to support having gender specific norms on the CVS.

Ethnicity

We ran one-way analysis of variance for the different ethnic groups in the current sample. The mean scores and standard deviations for each scale across the different ethnic groups are presented in Table 15, while the output from the ANOVA is reported in Table 16.

Table 15. Mean scores for ethnic groups

	Asian/Indian (*N* = 99)		Black (*N* = 426)		Coloured (*N* = 42)		White (*N* = 147)	
Scale	***M***	**SD**	***M***	**SD**	***M***	**SD**	***M***	**SD**
Service Orientation	29.51	3.37	29.86	3.61	28.88	3.61	28.65	3.18
Team Orientation	43.34	5.40	44.81	5.27	44.00	5.77	41.03	5.76
Influence	35.51	4.69	35.54	4.86	34.67	4.82	35.13	4.83
Creativity	32.80	4.42	34.00	4.29	32.21	4.35	30.90	4.18
Independence	36.85	4.98	37.97	5.04	38.12	4.64	35.53	4.90
Excitement	24.34	3.12	25.16	3.83	25.02	3.71	24.11	3.57
Career Development	38.77	3.97	38.63	3.59	38.26	3.86	38.21	3.80
Financial Rewards	43.57	4.95	43.22	6.43	45.24	5.24	41.98	5.95
Prestige	29.61	5.25	30.26	4.75	30.81	4.92	28.03	4.72
Security	26.66	3.90	27.08	3.64	27.52	3.39	25.56	3.65

Table 16. ANOVA results for ethnic group comparisons

Scale	*df*	*F*	*p*	n_p^2
Service Orientation	5, 2148	3.566	.003	.091
Team Orientation	5, 2148	10.372	.000	.155
Influence	5, 2148	.684	.636	.040

Creativity	5, 2148	21.198	.000	.222
Independence	5, 2148	13.917	.000	.180
Excitement	5, 2148	2.652	.021	.079
Career Development	5, 2148	.476	.794	.033
Financial Rewards	5, 2148	6.676	.000	.125
Prestige	5, 2148	5.036	.000	.108
Security	5, 2148	7.138	.000	.129

There were statistically significant differences between ethnic groups on all of the CVS scales, except Influence and Career Development. In order to determine where these differences were, we ran post-hoc tests, and these results are reported in Table 17. We only reported on those groups where statistically significant differences existed. Tukey HSD post-hoc results are reported for the scales that met the criteria for homogeneity of variance, while Games-Howell results are reported for those scales that did not meet these criteria.

Table 17. Post-hoc test results for ethnicity

								95% CI	
Scale	Post-hoc test	Ethnicity		Mean Diff	S.E.	Sig.	*d*	Lower Bound	Upper Bound
Service Orientation	Tukey HSD	Black	White	1.206	.343	.006	.36	.227	2.186
Team Orientation	Tukey HSD	Asian/Indian	White	2.316	.726	.018	.41	.246	4.386
		Black	White	3.780	.534	.000	.68	2.257	5.303
		Coloured	White	2.973	.977	.029	.52	.187	5.758
Creativity	Games-Howell	Asian/Indian	White	1.900	.563	.011	.44	.281	3.519
		Black	White	3.107	.403	.000	.73	1.951	4.263
Independence	Games-Howell	Black	White	2.441	.472	.000	.49	1.086	3.796
Excitement	Games-Howell	Black	White	1.048	.348	.034	.28	.049	2.048
Financial Rewards	Games-Howell	Coloured	White	3.259	.946	.012	.58	.491	6.026
Prestige	Tukey HSD	Black	White	2.229	.480	.000	.47	.861	3.597
		Coloured	White	2.776	.877	.020	.58	.273	5.278
Security	Tukey HSD	Black	White	1.513	.340	.000	.42	.542	2.484
		Coloured	White	1.959	.623	.021	.56	.183	3.735

Service Orientation

Black participants scored statistically significantly higher than White participants on the Service Orientation scale. The effect of this difference can be seen as moderate (0.36).

Team Orientation

White participants scored significantly lower on Team Orientation than all of the other ethnic groups, with moderate effects (0.41-0.68) for these differences. White participants seem to place less emphasis on Team Orientation as an important work value, than other ethnic groups in the current sample.

Creativity

Asian/Indian and Black participant groups both scored statistically significantly higher on the Creativity scale than the White group of participants in the current sample. The effects of these differences were moderate (0.44-0.73). This suggest that White participants place less value on Creativity as a work value, than the other two groups.

Independence

Black participants scored significantly higher than the White participants on the Independence scale, with a moderate effect size. This suggests that Black participants might place more value on Independence in the workplace than their white counterparts.

Excitement

On the Excitement scale white participants had statistically significant lower scores than the black participants. The effects of these differences were small (0.28), so there is insufficient evidence to support that the Black participants value Excitement more than white participants.

Financial Rewards

Coloured participants scored statistically significantly higher than white participants on the Financial Rewards scale, with a moderate effect (0.58). There were no significant differences between other ethnic groups on the Financial Rewards scale.

Prestige

Coloured and Black participants scored higher than white participants on the Prestige scale. The effects of these statistically significant differences were moderate (0.47-0.58), suggesting that White participants value Prestige lower than the other two ethnic groups.

Security

In the Security scale both Black and Coloured participants scored significantly higher than the White population group. The effect of the differences was moderate (0.42-0.56), and might indicate that Black and Coloured participants place more value on Security than their White counterparts.

Summary

White participants scored lower than all other ethnic groups across all of the CVS scales. In several scales White participants scored statistically significantly lower than one, or more, of the other population groups. The results from the ethnic group comparisons might indicate a need for ethnicity specific norm groups.

FACTOR ANALYSIS

In order to establish construct validity, we conducted an exploratory and confirmatory factor analysis to determine if the factor structure of the CVS matches that of the theoretical model. Construct validity is an estimate of how well the items within a test represents the construct (concept) that the test aims to measure and can be determined in a number of ways, including factor analysis and item analysis (see the section on Rasch analysis).

Exploratory factor analysis

We conducted an exploratory factor analysis using Principal Axis Factoring Extraction and Direct Oblimin Rotation with 200 iterations, specifying 10 factors for extraction. These 10 factors explained 39.41% of the total variance in the CVS dataset. The pattern matrix for the 10 factors are presented in Table 18.

Table 18. CVS factor loadings

	Factor									
	1	2	3	4	5	6	7	8	9	10
SO1	.052	-.009	.061	.175	.052	-.069	-.079	-.168	**.401**	.055
SO2	.072	-.018	.089	.147	-.027	-.019	-.043	-.117	**.536**	.156
SO3	**.264**	.050	.026	-.026	.231	-.058	-.049	-.089	**.298**	.130
SO4	.049	-.025	-.008	.056	.095	.182	-.024	.040	**.519**	-.007
SO5	.192	.113	.056	-.004	.181	-.081	-.074	-.093	**.340**	.073
SO6	-.002	-.003	.041	.103	-.007	.087	-.001	.045	**.643**	-.061
SO7	.157	-.126	-.021	.100	.059	.129	.068	-.017	**.605**	-.135
TO1	-.032	-.011	.011	**.558**	.113	-.039	-.149	-.037	.103	.203
TO2	-.065	.054	.007	**.627**	-.028	.037	.012	.123	-.090	-.083
TO3	.076	-.012	.016	**.474**	.110	-.029	-.075	-.014	.131	.246
TO4	.018	.086	-.092	**.486**	.021	-.026	.040	.003	.161	-.025
TO5	.183	.038	.046	.169	.000	.046	**.281**	.015	.082	.040
TO6	.014	-.035	.042	**.637**	-.019	-.005	-.057	.109	.126	.137
TO7	.217	.031	.082	**.318**	-.028	.146	.053	-.047	.074	.009
TO8	-.027	.001	-.038	**.713**	.097	-.019	.003	-.044	.083	-.155
TO9	.053	.002	.225	**.253**	-.010	-.087	**.292**	-.012	-.038	-.147
TO10	-.104	.154	.085	**.267**	.221	-.049	.000	-.081	.217	.022
TO11	.228	.044	-.042	.132	.043	-.105	**.293**	.127	.039	-.018
TO12	-.065	-.072	-.013	**.651**	-.010	.004	.043	.057	.153	.017
INF1	-.053	.024	-.063	.063	**.531**	.053	-.092	-.035	.128	.075
INF2	.068	.084	.090	.040	**.493**	.116	-.045	-.050	.060	.057
INF3	.133	.121	.048	-.039	**.406**	-.035	.090	-.009	.055	.021
INF4	-.115	.000	.051	.074	**.521**	.057	-.050	.161	.110	.083
INF5	.097	-.015	**.303**	.080	**.378**	.043	-.026	.054	.077	.131
INF6	-.042	-.003	.041	.073	**.601**	.049	-.007	.201	.055	-.093
INF7	.123	-.014	.120	.008	**.391**	-.151	.042	-.022	.114	.172
INF8	.020	.040	.128	.148	**.537**	-.077	.034	-.015	.049	.081
INF9	.046	-.058	-.072	.066	**.446**	.151	.014	.140	.183	-.039
CRE1	.147	.049	**.279**	.003	.126	.198	-.058	-.087	-.014	**.290**
CRE2	.153	.062	**.543**	.112	.016	.034	.010	-.127	-.047	.054
CRE3	.166	-.018	**.489**	.044	.070	.122	.011	.035	-.022	**.255**
CRE4	.203	-.029	**.458**	-.028	.161	.107	-.046	-.007	.071	**.273**
CRE5	**.256**	-.040	**.495**	.109	-.036	.143	.055	-.031	.012	.081
CRE6	.134	-.132	**.607**	.013	.068	.070	-.030	.170	.004	.149

CRE7	.233	-.056	**.530**	.099	-.064	.024	.063	.042	.021	.042
CRE8	.124	-.076	**.497**	-.032	.101	-.006	.097	.158	.063	.048
IND1	.067	.023	-.115	-.080	**.337**	.196	.182	-.143	-.041	.045
IND2	.166	.079	.226	-.014	.195	-.130	.190	-.079	.066	.088
IND3	-.141	.006	.056	.022	-.109	.214	**.395**	.021	-.065	-.044
IND4	-.056	.084	.000	-.077	.003	.077	**.559**	-.042	-.004	-.076
IND5	-.037	-.006	.122	-.045	.189	-.071	**.560**	-.013	-.044	.118
IND6	-.117	.095	**.498**	-.054	-.016	-.018	.148	.058	.123	-.191
IND7	-.037	.101	.247	-.055	.015	-.024	**.447**	.021	-.002	-.172
IND8	.040	.014	.227	.011	**.278**	-.106	**.390**	.028	-.039	.063
IND9	.075	.110	.103	-.003	.176	.011	**.375**	-.048	.035	-.235
IND10	.008	.076	-.083	.189	-.171	.008	**.407**	.056	-.021	.031
EXC1	.201	.099	.024	**.257**	.057	.243	.042	-.218	-.093	-.041
EXC2	-.065	.001	.102	.065	.109	**.528**	-.014	-.044	.063	.102
EXC3	-.022	.182	**.281**	.052	-.001	.199	.057	.026	.088	.040
EXC4	.117	-.072	.116	.232	-.123	**.300**	**.269**	.008	-.006	.020
EXC5	-.085	-.012	.106	.007	.029	**.582**	-.012	.073	.157	-.011
EXC6	.148	-.046	.189	.177	-.170	**.302**	**.297**	.059	.024	.034
EXC7	-.010	-.030	.034	.012	**.262**	.237	.059	.139	.059	**.269**
CAD1	.099	.039	-.144	-.074	.047	.204	-.026	.021	-.020	-.017
CAD2	**.497**	.093	.119	.044	-.057	.056	-.021	.026	.078	.181
CAD3	**.571**	.099	.081	.007	-.011	-.022	-.093	-.030	.104	.062
CAD4	**.615**	.018	.133	.036	-.067	.007	-.023	.010	.110	.091
CAD5	**.685**	.053	.065	.007	-.052	-.121	-.036	.027	.076	.039
CAD6	**.569**	-.078	.135	.003	.025	.009	.006	.067	.134	.084
CAD7	**.577**	.065	-.030	.024	.157	.096	-.016	.051	.052	-.057
CAD8	**.374**	-.044	-.021	.041	.225	.217	-.002	-.022	.090	.066
CAD9	**.634**	.029	.067	-.010	-.016	-.006	-.010	.042	.140	-.015
FIR1	-.027	**.666**	.045	.035	.003	-.008	-.058	-.013	-.106	.027
FIR2	.076	**.692**	-.016	.014	-.012	-.055	-.041	-.095	.030	-.016
FIR3	-.108	**.527**	.046	-.040	-.066	.188	.062	.207	-.009	-.227
FIR4	.157	**.278**	.028	-.045	.134	.111	.063	.007	.010	.153
FIR5	.020	**.457**	-.014	-.006	.105	.119	.119	.169	-.024	.110
FIR6	-.066	.154	**.608**	-.026	.027	-.021	.029	.050	.083	-.150
FIR7	-.045	**.671**	.137	-.060	-.027	.065	-.011	.185	-.042	-.062
FIR8	-.064	**.357**	.037	.042	.064	.002	.147	.150	-.032	-.073
FIR9	-.090	**.346**	.092	.051	**.347**	.018	-.061	.127	-.053	.162
FIR10	.096	**.477**	.058	-.023	.115	.057	.110	.161	-.070	.001
FIR11	-.124	**.382**	.083	-.036	-.019	.127	.054	**.352**	.017	-.186
PR1	.181	.131	-.045	.082	.158	.055	-.024	**.323**	-.130	-.058
PR2	.197	-.025	.044	.241	.217	-.062	-.060	.192	.124	.004
PR3	-.110	.108	.018	.057	.095	.153	.053	**.470**	-.094	-.076
PR4	.064	.067	.117	.094	-.007	-.052	.018	**.611**	-.004	.079
PR5	.075	.141	.124	.065	.006	-.014	-.004	**.597**	.025	-.034
PR6	.193	.001	.042	.172	**.383**	.049	-.025	.157	.067	-.219
PR7	**.295**	.016	.008	-.013	**.280**	-.008	.026	**.424**	-.029	-.086
PR8	.027	.211	.015	.072	.030	-.015	.000	**.481**	.006	.164
SEC1	.190	**.361**	-.090	.093	.014	-.156	.041	-.108	.144	.031
SEC2	-.112	.070	-.041	-.074	-.089	-.147	**.337**	.165	.118	.091
SEC3	.120	**.326**	-.153	.015	-.098	-.219	.149	.155	.169	.164
SEC4	.112	.105	-.079	-.001	.098	-.165	.176	.122	.134	.189
SEC5	.074	**.438**	-.169	.064	-.085	-.079	.102	.042	.113	.013
SEC6	.193	.204	.034	.027	-.043	-.185	.056	**.273**	.173	.145
SEC7	-.226	.009	-.143	-.021	-.100	-.059	**.338**	**.281**	.096	.080

In Table 19 we provide a simplified factor loading structure.

Table 19. Simplified factor loading structure from the exploratory factor analysis

1	2	3	4	5	6	7	8	9	10
CAD2	FIR1	CRE2	EXC1	FIR9	CAD1	IND3	PR1	SO1	CRE1
CAD3	FIR2	CRE3	PR2	IND1	EXC2	IND4	PR3	SO2	EXC7
CAD4	FIR3	CRE4	TO1	INF1	EXC4	IND5	PR4	SO4	SEC4
CAD5	FIR4	CRE5	TO2	INF2	EXC5	IND7	PR5	SO5	
CAD6	FIR5	CRE6	TO3	INF3	EXC6	IND8	PR7	SO6	
CAD7	FIR7	CRE7	TO4	INF4		IND9	PR8	SO7	
CAD8	FIR8	CRE8	TO6	INF5		IND10	SEC6		
CAD9	FIR10	EXC3	TO7	INF6		SEC2			
SO3	FIR11	FIR6	TO8	INF7		SEC7			
	SEC1	IND2	TO10	INF8		TO11			
	SEC3	IND6	TO12	INF9		TO5			
	SEC5			PR6					

Although some of the items loaded on different factors than expected, nine of the CVS factors can be clearly identified from Table 19:

1. Factor 1 – Career Development
2. Factor 2 – Financial Rewards
3. Factor 3 – Creativity
4. Factor 4 – Team Orientation
5. Factor 5 – Influence
6. Factor 6 – Excitement
7. Factor 7 – Independence
8. Factor 8 – Prestige
9. Factor 9 – Service Orientation

The only scale that didn't have the expected factor loadings, was the Security scale with items from this scale loading mostly unto the Financial Rewards and Independence scales. Items that did not load as expected will be further investigated.

Item Parcelling

We also ran an exploratory factor analysis with item parcels (De Bruin, 2004) to investigate whether the theoretical structure of the CVS would be represented in the current data.

Item parcelling as an analytical tool within factor analysis has been debated by many scholars in recent years, but with appropriate use provides reliable and valid indicators of the constructs under investigation (Little, Rhemtulla, Gibson, & Schoeman, 2013). In order to do this, we parcelled items (ranging from 2–4 items per parcel) together within each one of the CVS scales in order to accommodate for items that did not clearly load unto their expected factors.

Table 20 provides the factor loadings based on an exploratory factor analysis with a Principal Axis Factoring Extraction method and a Direct Oblimin Rotation applied, with 200 iterations and 10 factors specified for extraction.

Table 20. CVS factor loadings with item parcels

	Factor									
	1	2	3	4	5	6	7	8	9	10
PSO1	.231	-.088	.097	.245	.081	-.179	-.064	-.064	.177	**-.340**
PSO2	.234	.073	.008	.061	-.004	-.153	-.060	.040	.053	**-.505**
PSO3	.072	.001	.022	.187	-.029	-.064	.007	.084	.015	**-.584**
PTO1	.099	-.040	.010	**.802**	.027	.029	-.087	-.005	-.028	.039
PTO2	-.062	.008	.060	**.647**	.000	-.053	.066	.033	-.054	-.144
PTO3	-.113	.018	-.014	**.564**	.080	.011	.145	.122	-.015	-.115
PINF1	**.762**	.036	.002	.011	-.008	-.065	.061	.003	.055	-.007
PINF2	**.668**	.012	.002	.010	.137	.093	.020	.085	-.186	-.097
PINF3	**.539**	-.040	.041	.029	.090	-.075	.064	.009	-.107	-.180
PCRE1	.145	.027	-.022	.110	**.652**	-.136	.006	.053	.132	.130
PCRE2	.078	-.010	-.028	.023	**.783**	-.085	-.034	.084	-.020	-.020
PCRE3	-.034	.004	.036	.034	**.659**	-.040	.091	.035	-.069	-.059
PIND1	.110	-.002	-.018	-.016	-.083	-.063	**.669**	.061	.068	.080
PIND2	-.037	.105	.018	-.023	.226	.134	**.582**	-.014	-.053	-.063
PIND3	-.047	.025	.115	.083	.024	-.027	**.597**	.001	-.072	-.050
PEXC1	.121	.110	-.065	.153	.085	-.092	.037	**.467**	.144	.117
PEXC2	-.057	-.004	-.006	.005	-.042	.021	.001	**.861**	-.008	-.044
PEXC3	.056	-.018	.095	-.021	.150	-.060	.062	**.453**	-.076	-.058
PCAD1	.011	.054	-.028	-.018	-.006	**-.616**	-.006	.018	.006	.032
PCAD2	-.100	-.095	.074	.047	.198	**-.756**	.013	-.050	-.052	-.087
PCAD3	.078	-.041	.003	-.015	.009	**-.693**	.041	.106	-.072	-.107
PFIR1	-.041	**.765**	.064	.033	-.101	-.021	.031	.001	.053	.050
PFIR2	.038	**.501**	-.014	-.019	.275	-.053	.131	.012	-.014	-.090
PFIR3	.103	**.704**	.094	-.003	.056	.041	-.022	.016	-.040	.007
PFIR4	-.051	**.707**	.004	-.060	-.004	.020	.056	.046	-.180	-.054
PPR1	.170	.200	-.006	.207	-.060	-.092	.047	.022	**-.392**	.126
PPR2	.071	.177	.117	.164	.064	-.046	.016	.045	**-.503**	.004
PPR3	.131	.164	.171	.002	.020	-.179	-.005	.004	**-.495**	.032
PSEC1	-.009	.127	**.556**	.071	-.052	-.041	.096	-.112	.085	-.026
PSEC2	.049	.110	**.755**	.026	.034	-.102	-.089	.001	.045	-.025
PSEC3	-.023	-.081	**.784**	-.042	-.003	.097	.048	.081	-.108	.035

The results obtained from the factor analysis conducted on item parcels yielded a clearer factor structure, where all of the factor parcels loaded as expected.

Confirmatory factor analysis

In order to determine whether the factor structure of the current dataset matches the model, we ran a confirmatory factor analysis (CFA), with a Weighted Least Square Means estimator (WLSM) on R using the Lavaan package (Yves, 2012). This allowed us to compare the current dataset's structure against a statistical model. The results from the CFA are presented in Table 21. We decided to run the CFA on both the item level data and the item parcels.

Table 21. CFA fit measures

Fit indices:	Item level	Parcel level
Comparative fit index (CFI)	.816	.927
Tucker-Lewis index (TLI)	.810	.913
Robust CFI	**.875**	**.970**
Robust TLI	**.870**	**.964**
Root mean square error of approximation (RMSEA)	.075	.077
90 % confidence interval	.074-.075	.074-.080
p-value RMSEA	.000	.000
Robust RMSEA	**.062**	**.049**
90% confidence interval	**.061-.062**	**.048-.051**
Standardized Root Mean Square Residual (SRMR)	**.067**	**.048**

The results from the CFA indicated that the item level factor structure of the CVS with the current sample did not fit the statistical model. Literature suggests that CFI and TLI fit statistics should be above 0.95 (Hu & Bentler, 1999), while some authors suggest 0.90 as sufficient. The results of the CFA of item parcels showed good model fit, with the CFI and TLI fit indices being above 0.95. Hu and Bentler (1990) suggest that the RMSEA value should be below 0.06. Therefore, the item parcel model fit the statistical model better than an item level CFA model. Other authors suggest that RMSEA values up to 0.08 are acceptable (Marsh, Hau & Wen, 2004). Literature further suggests that SRMR values below 0.10 (some authors prefer 0.08, see Hu & Bentler, 1999) are indicative of good model fit (Kline, 2005), suggesting that both the item level and the item parcel level models comply with the SRMR cut-offs.

Overall, the item parcel model fitted the statistical model better, and the results suggest that the factor structure of the CVS for the current sample conforms with the theoretical structure of the CVS as proposed in the original CVS manual (McNab, et al., 2005).

RASCH ANALYSIS[2]

The Rasch model is a statistical model that allows for the analysis of categorical data, such as responses to questions in psychometric assessments on two interactive levels, namely the ability or trait of the respondent, and the level of item difficulty (Rasch, 1960). In other words, a respondent's response to any given question is a logistic function of the differences between their ability to respond to/tendency to endorse the item, and the difficulty of the item (Green & Frantom, 2002). The unit of measurement in Rasch analysis is the logit (or log-odds unit). The mean logit score is set at 0.0, with higher scores indicating greater difficulty, and negative scores indicating lesser difficulty (Bond & Fox, 2007).

This section will address each scale of the CVS separately, and will address item fit and differential item functioning for each one of the scales.

Item Analysis

Item fit to the Rasch model is an indication of how well responses from candidates conform to a logical pattern (Green & Frantom, 2002). The items can 'underfit' (be too unpredictable, random), or 'overfit' (be too predictable) the model. Based on these fit indices, items can be flagged for potential rewrite or removal from the assessment. The specific fit indices we looked at are:

1. **Measure** – this statistic refers to the likelihood that a participant will endorse an item. Items with positive values are more difficult for the participants to endorse, while items with a negative value are easier to endorse.
2. **Mean-square statistic (MNSQ)** – this statistic gives an indication of how well each item fits the predictions of the Rasch model. These values are expected to be close to 1.0. Items with good fit will generally have scores ranging between 0.70 and 1.35 logits (Linacre, 2015). Items with a score lower than 0.70 might be redundant, i.e., they are measuring the same thing as other items and do not add any additional information. Items with a score greater than 1.35 might be measuring a different construct to what the test is intended to measure.
3. **Standardised fit statistics (ZSTD)** – this is a z-score output of a t-test to determine how well the data fit the Rasch model. Scores should be central around 0.0. Scores higher than 0.0 indicate a lack of predictability – the item did not function as the Rasch model predicted, while scores below 0.0 indicate too much predictability – i.e., not enough variance in response patterns.

Differential Item Functioning (DIF)

Differential item functioning refers to specific items within psychometric assessments that cause test takers from various subgroups, i.e., ethnic, gender, etc., to score differently despite having equal ability to perform on the assessment (Westers & Kelderman, 1991). Therefore, psychometric assessments containing items with significant DIF might be unfair or biased to specific population subgroups, in direct contradiction to the Employment Equity Act (No 55 of 1998), Chapter 2, Paragraph 8, which state that:

> "*Psychometric testing and other similar assessments of an employee are prohibited unless the test or assessment being used –*

[2] For a more detailed explanation and tables containing the statistical output from the Rasch Analysis, please see *CVS SA Item Analysis Technical Report.*

(a) has been scientifically shown to be valid and reliable;
(b) can be applied fairly to employees; and
(c) is not biased against any employee or group."

In order to investigate if potential DIF existed in any of the CVS items, we ran DIF analyses across the various sample subgroups, including age (excluding the 51–60-year-old age groups due to sample size), gender, and ethnicity (only black and white participants).

For the interpretation of DIF in the current study, we looked at the size of the DIF contrast between the reference group and the focal group, along with the probability statistic from the Mantel Chi-square calculations. The Mantel calculation is less sensitive to missing data than the Rasch-Welch statistic, which is also reported. Literature suggests than any item with a DIF contrast greater than 0.5 logits should be identified for potential DIF, and the probability statistic will be indicative of the statistical significance of this difference.

We provide an overview of the fit statistics for the 11 scales of the CVS in Table 22. From the table we can see that most of the CVS scales showed overall good fit statistics. We also see that the person and item measures show good differentiation between ability and endorsability of the items.

Table 22. Fit statistics across the CVS scales

Scale	Person		Item		Mean	
	Real Sep	Rel	Real Sep	Rel	MNSQ	ZSTD
Service Orientation	1.66	.73	17.37	1.00	1.03	.50
Team Orientation	1.95	.79	8.26	.99	1.00	-.40
Influence	1.99	.80	18.24	1.00	1.01	.10
Creativity	2.03	.80	14.81	1.00	.99	-.50
Independence	1.47	.68	23.16	1.00	.99	-.50
Excitement	1.43	.67	21.32	1.00	.99	-.40
Career Development	1.53	.70	23.04	1.00	.99	-3.6
Financial Rewards	1.91	.78	23.79	1.00	1.00	-.90
Prestige	1.73	.75	24.17	1.00	.99	.10
Security	1.32	.63	34.86	1.00	1.00	.10

Service Orientation

Item Fit

Most of the items from the Service Orientation scale showed good item fit and fell well within the accepted ranges for their mean score statistics. Only item SO5 had an infit MNSQ statistic above 1.35, indicating that this item might be measuring a construct that is not related to the others. While conducting the factor analyses, this item did load unto the same scale as other Service Orientated items (0.340). This item will be flagged and investigated further within the differential item functioning section.

Figure 2 represents the item-person map for the current sample on the Service Orientation scale. From the map it can clearly be seen that there is limited spread across the items in terms of item endurability, with the majority of the sample being most likely to endorse the items in a positive manner, i.e., rating of 4 and higher on the Likert-type scale.

This could suggest that the current items aren't effectively differentiating between participants, or that the current sample rated themselves higher on Service Orientation due to the contexts of their various assessment situations. It is expected that participants will rate themselves higher on self-report scales.

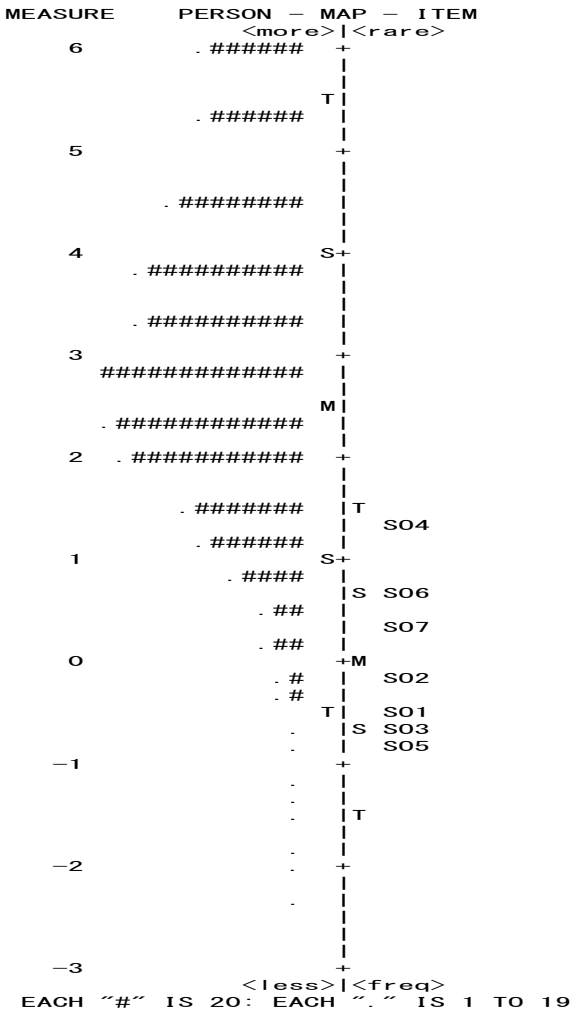

Figure 2. Item map: Service Orientation

Differential Item Functioning

Age

The DIF plot for the five age groups for the Service Orientation scale is presented in Figure 3. Several items seem to function differently for the different age groups, but the biggest differences seem to be between the youngest age group (15–20) and the other age groups. Apart from items SO1, SO2, and SO3, the rest of the DIF plot seems to follow a similar pattern for all of the age groups. This could suggest that the differences are due to actual differences between the groups and not due to DIF. We will rely on the results from the DIF statistics to determine whether specific items need to be flagged for potentially causing DIF.

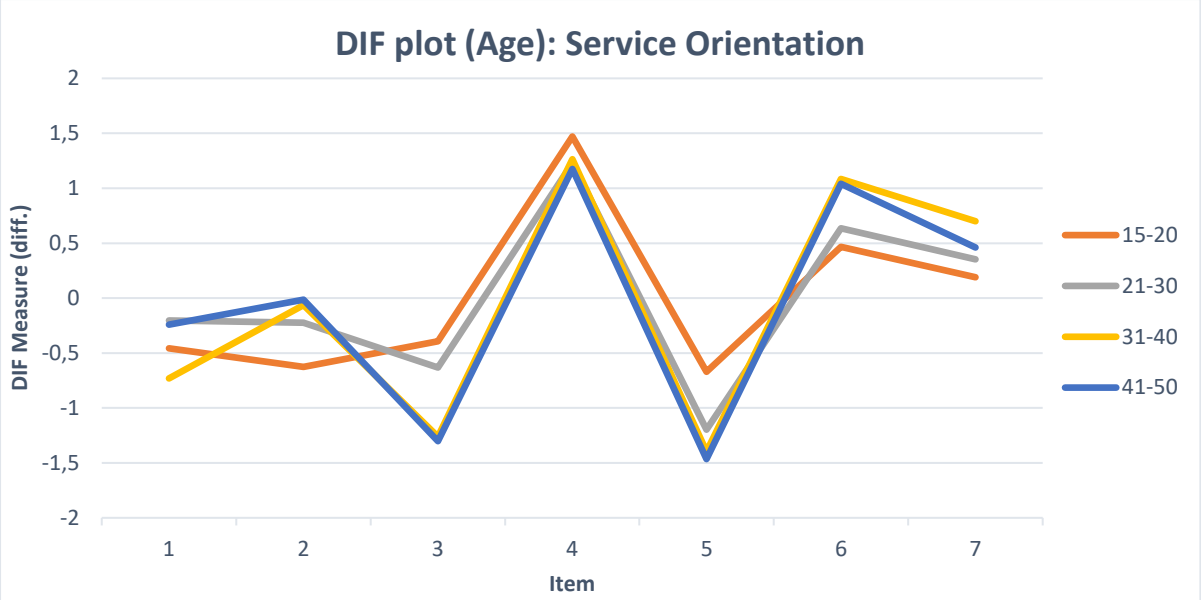

Figure 3. DIF plot (Age): Service Orientation

Item SO5 was flagged for potentially causing DIF between the 15–20-year-old group and three other age groups (21–30, 31–40, and 41–50). This item seemed to be more difficult to endorse for the youngest age group in the sample. Items SO2 and SO3 appear to be causing DIF between the youngest age group and all of the other age groups, except the 21–30-year-olds. Item SO2 is a lot easier to endorse for the 15–20-year-olds, while item SO3 was more difficult to endorse. Item SO6 was more difficult for the 15–20-year-old participants than for the 31–40 and 41–50-year-old groups. The 15–20-year-old participants found it significantly easier to endorse item SO7 than the 31–40-year-old group. The probability of this item causing DIF in other samples with age differences was statistically significant, and the item needs to be further investigated to understand what is causing the DIF.

Overall it seems as if the youngest age group is responding to the items related to Service Orientation in a different manner than participants who might have more experience in working industries. This might indicate that the scale is not appropriate for use with participants who are not from the working adult population.

Item SO3 was flagged for potentially causing DIF between the 21–30-year-old group and the 31–40 and 41–50-year-old group. For both of these comparisons there was a statistically significant probability that this item would also cause DIF in other samples between these age groups. The results of the DIF analysis also suggest that item SO1 is potentially causing DIF between participants from the 21–30 and the 31–40-year-old population groups. The significant probability that the Mantel-Haenszel test yielded indicates that this item will most likely also cause DIF in other samples between these age groups. It seems that the 31–40-year-olds found this item easier to endorse than their younger counterparts.

None of the items were flagged for potentially causing DIF between the 31–40-year-old and 41–50-year-old candidates.

Gender

Figure 4 provides the graphical representation of DIF between men and women on the Service Orientation scale. From the graph it can be seen that the DIF measure follows a unidirectional pattern for both genders, with only item SO5 appearing to have a large DIF contrast for men and women. Only item SO5 was identified for potentially causing DIF between different genders on the Service Orientation scale. The Mantel Chi-square test yielded a statistically significant result, suggesting that item SO5 is probable to cause DIF between gender groups in other samples as well. Men found it more difficult to endorse this item than women. It is important to note that these differences had small effect sizes, and that the probability that real measurable differences between how men and women respond, could be restricted to the current sample only.

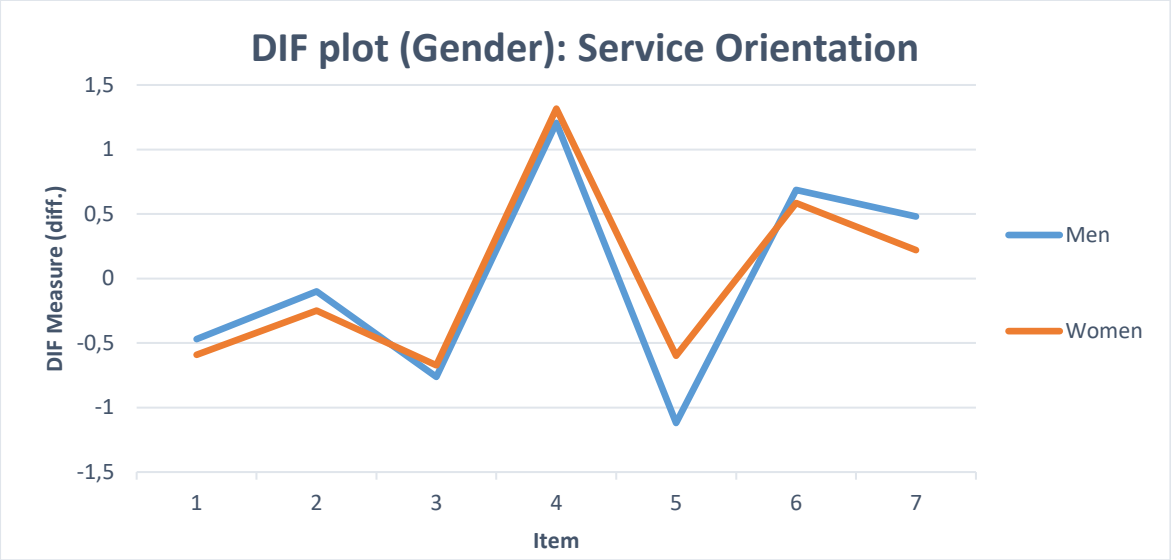

Figure 4. DIF plot (Gender): Service Orientation

Ethnicity

The DIF plot for the black and white participants on the Service Orientation scale is presented in Figure 5. From the graph it is clear that the two groups had a unidirectional performance curve, except for item SO2.

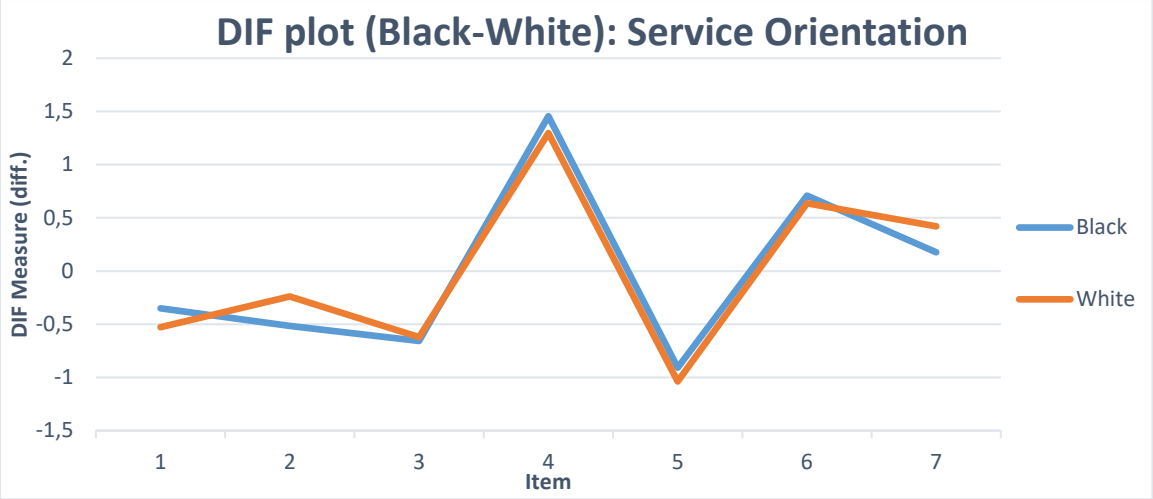

Figure 5. DIF plot (Black-White): Service Orientation

None of the items were flagged for potential DIF between black and white participants on the Service Orientation scale.

Team Orientation

Item Fit

Upon investigating the results from the mean square statistic for the items in the Team Orientation scale, two items were flagged for potential non-fit. Item TO6 fell below the 0.70 range, suggesting that this item might be redundant and not adding any additional information to the scale. Item TO10 yielded an outfit mean square statistic above the 1.35 range, while having an infit statistic of 1.35, suggesting that this item is potentially measuring a different construct than Team Orientation. When conducting the factor analysis, this item loaded weakly (<0.300) unto the Team Orientation scale, suggesting that this item might need to be removed or adjusted.

The positioning of the item difficulty against the participants' ability for the Team Orientation scale is presented in Figure 6. From the item map it can be seen that most of the participants from the current sample would found the items easy to endorse or agree with, in other words most of the participants rated the value statements as important or very important on the Likert-type scale. This would suggest that the items from this scale was not sufficient in eliciting less agreeable responses from the current sample, and could indicate that the items aren't effectively measuring a true reflection of how much participants value Team Orientation. It could also be indicative that the majority of the participants did value Team Orientation, and the results might look different in a different sample.

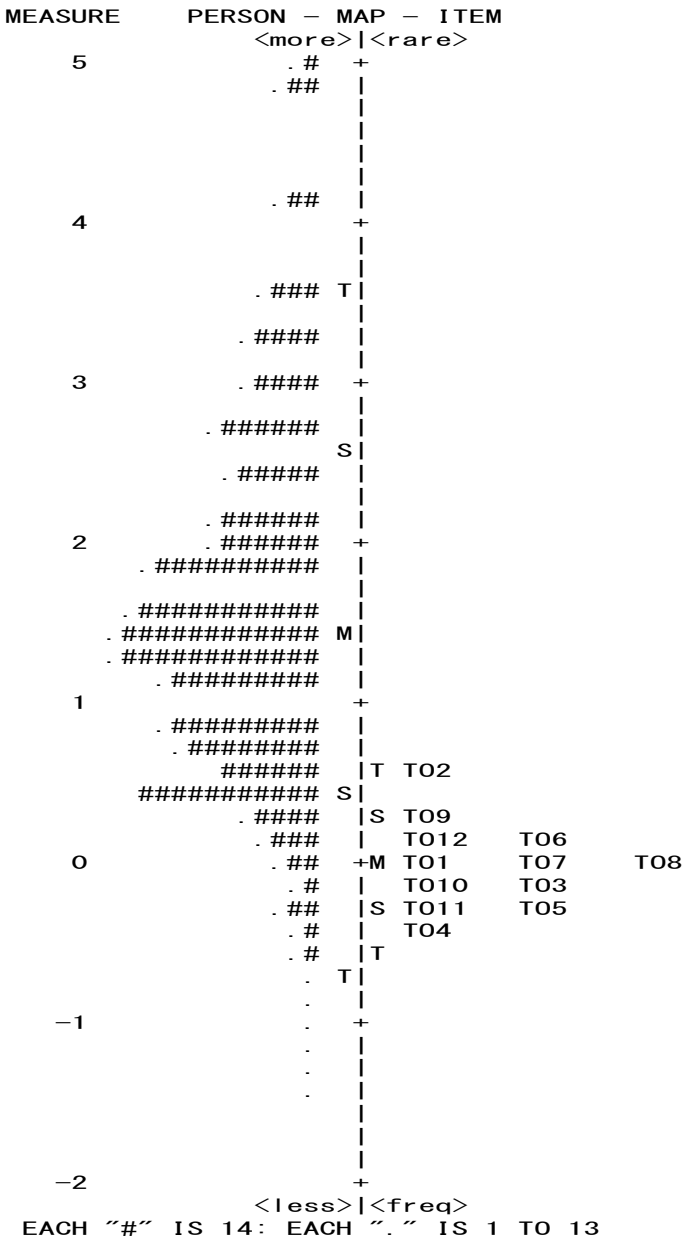

Figure 6. Item map: Team Orientation

Differential Item Functioning

Age

The Team Orientation DIF plot for the different age groups are presented in Figure 7. From the plot it is clear that the DIF plot does not follow a unidirectional pattern for all of the age groups, whereas the plots for the two younger age groups seem to follow a more unidirectional pattern, while the plots for the two older age groups seem to also follow a more unidirectional pattern.

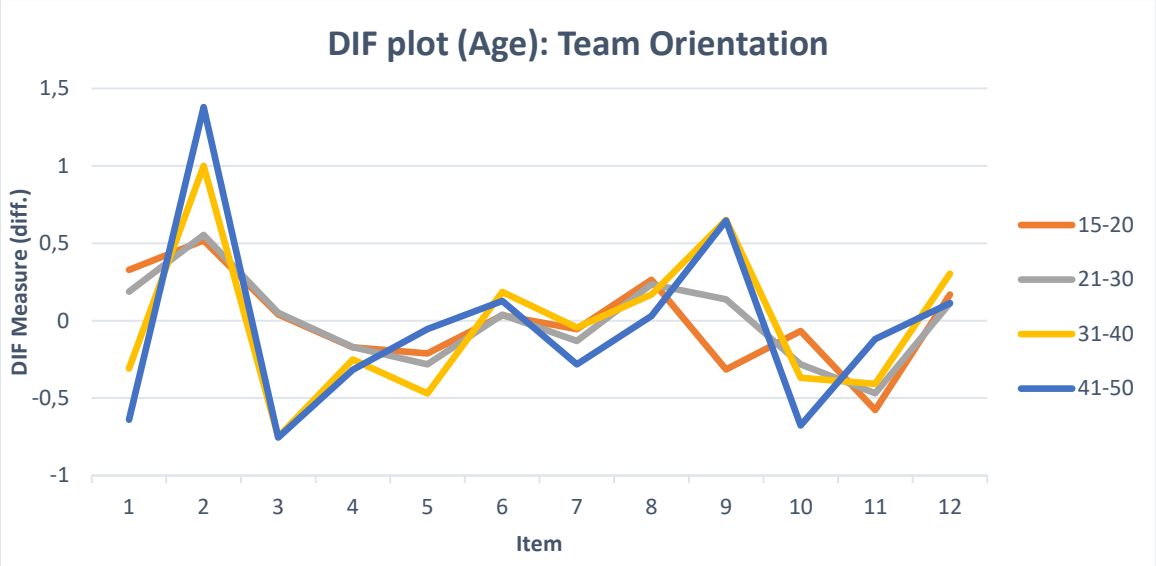

Figure 7. DIF plot (Age): Team Orientation

Item TO9 was flagged as causing DIF between the 15–20-year-old age group, and all of the other age groups in the analysis. The probability of this item causing age-related DIF in other samples was found to be statistically significant. It seems that it is easier for the youngest population group to endorse this item than for any of the other population groups. Items TO1, TO2 and TO3 were all identified as potentially causing DIF between the 15–20-year-olds, and the 31–40 and 41–50-year-old groups. It seems as if the younger participants (15–20) found it more difficult to endorse items TO1 and TO3. Item TO2, which was easier for the younger population group to endorse than for the older groups were generally more related to sociability. The impact of the DIF for all three of these items was statistically significant, indicating that these items will potentially cause DIF in other samples as well as between these age groups. It might be that these are actual differences in the aspects of Team Orientation that different age groups value, and could be influenced by the level of exposure participants have had to a working environment. When analysing for DIF between the 15–20-year-old group and the 41–50-year-old group, a further two items were identified for potential age-related DIF: items TO10 and TO11. Younger participants found item TO10 easier to endorse than the older population group, while the older population group found it easier to endorse item TO11 than the 15–20-year-olds. Both of these items yielded statistically significant Mantel-Haenszel statistics, suggesting that this pattern could also emerge in future samples. It is important to note that the effect size of the DIF for item TO11 between these two age groups did not exceed the 0.5 logit estimate, indicating that the differences are perhaps only a result of the current sample.

Next, we provide an overview of the DIF statistics for the 21–30-year-old group as the reference group, with the 31–40 and 41–50-year-old groups as the focal groups. Items TO1, TO2 and TO3 were all identified as potentially causing DIF in both age group comparisons, with the younger population group (21–30-year-olds) finding it easier to endorse item TO2, and more difficult to endorse items TO1 and TO3. The impact of the DIF for all three items were statistically significant across both age group comparisons, indicating that they could also cause DIF in future samples between these age groups. Item TO9 was also flagged as potentially causing DIF between the 21–30-year-old and the 31–40-year-old groups. The Mantel-Haenszel test resulted in a statistically significant probability that this item might cause DIF in other samples as well.

The 31–40-year-olds found it easier to endorse items TO2 and TO5, and the Mantel-Haenszel test revealed statistically significant probabilities that their items will cause age-specific DIF in these two groups in following samples. The DIF contrasts for these items did not exceed the 0.5 logit estimate to be identified as causing significant DIF, and it is argued that the statistically significant results obtained from the Mantel-Haenszel test are influenced by factors relating to the current sample.

Gender

The DIF plot for gender DIF for the Team Orientation scale is presented in Figure 8.

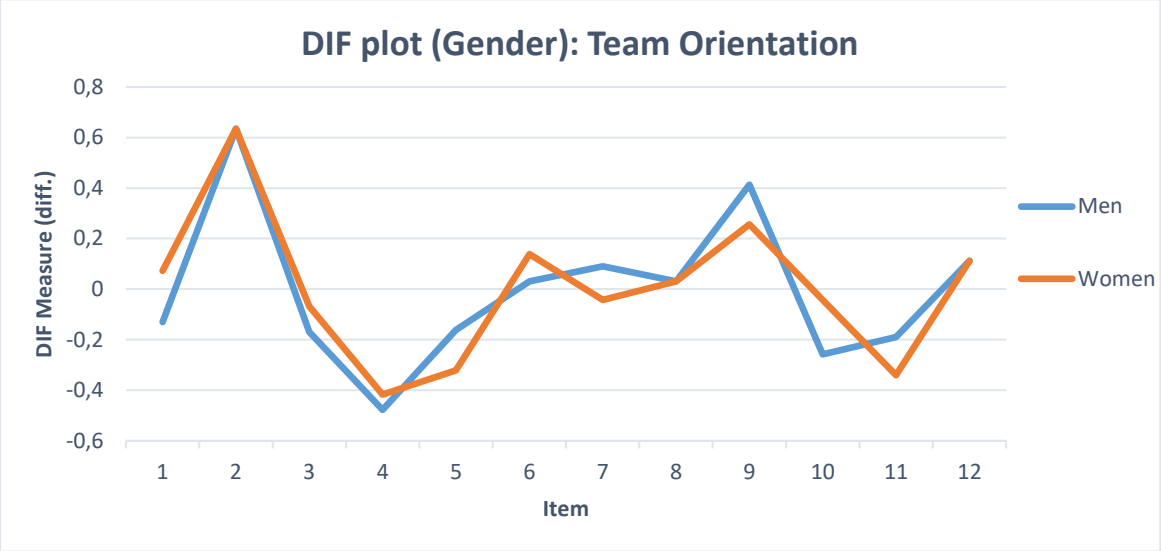

Figure 8. DIF plot (Gender): Team Orientation

The DIF plot for gender follows a unidirectional pattern up until around item TO5 where we start to see deviations between the two plots for several items. Five items were identified with statistically significant Mantel-Haenszel test results for DIF. Men found it easier to endorse items TO1, TO6 and TO10. Women found items TO5, TO9 and TO11 easier to endorse. Although all of these items yielded statistically significant results, it is important to note that the DIF contrasts for none of them exceeded the 0.5 logit estimate, suggesting that the effect of the significance is small and could only be restricted to the current sample.

Ethnicity

A graphical representation of the DIF measurement statistics for the two ethnic groups are given in Figure 9. When investigating the DIF plot, we can see that the two ethnic groups do not follow a unidirectional pattern in terms of DIF measure statistics. There are several items that appear to be easier or more difficult to endorse for each of the ethnic groups.

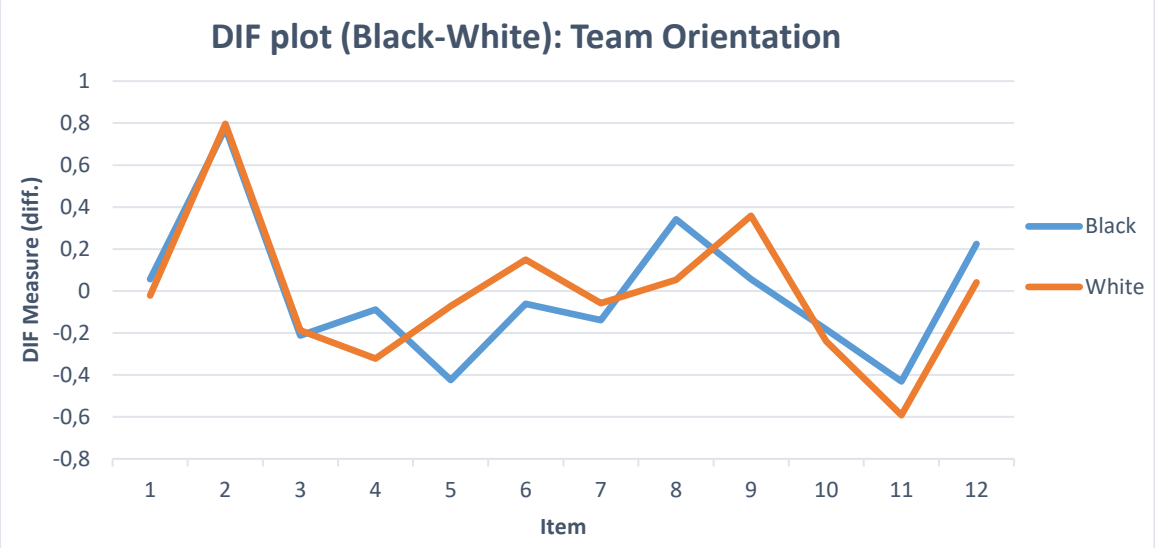

Figure 9. DIF plot (Black-White): Team Orientation

Four items from the Team Orientation scale were identified with statistically significant Mantel-Haenszel test results, i.e., likely to cause DIF between black and white participants in other samples. The effect of the DIF is however small (<0.5 logits) and it is therefore arguable that these differences in response patterns for black and white candidates are influenced by the specific sample.

Influence

Item Fit

The statistics for measuring difficulty/endorsability showed a relatively equal spread between items that were easy to endorse and items that were more difficult to endorse. Based on the mean square statistics for both the infit and outfit analyses, none of the items from the Influence scale were flagged for potential misfit.

When investigating the item map (Figure 10), it is clear that the majority of the participants from the current sample found the items from the Influence scale easy to endorse, i.e., gave ratings of 4 and higher on the 5-point Likert-type scale.

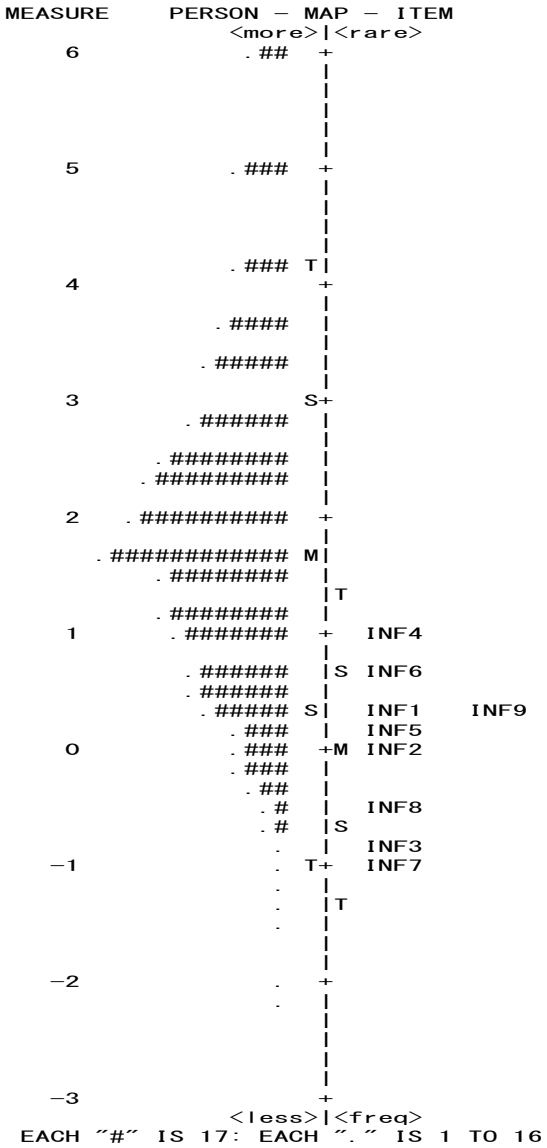

Figure 10. Item map: Influence

Differential Item Functioning

Age

We plotted the DIF measures for the items of the Influence scale for the four distinct age groups in Figure 11.

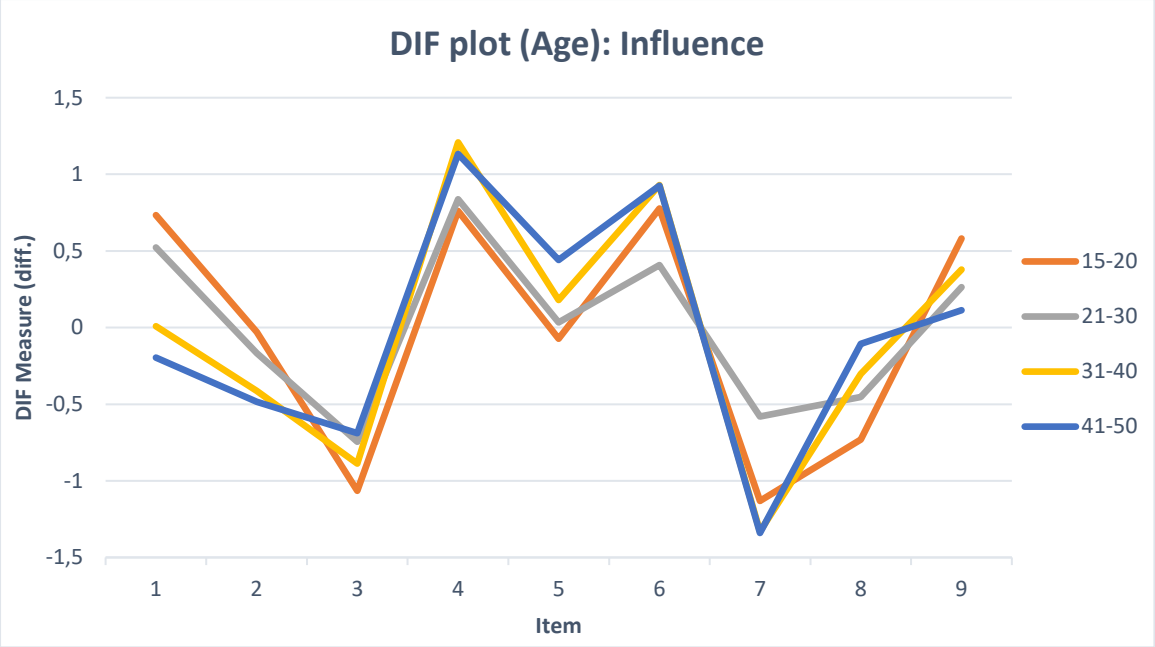

Figure 11. DIF plot (Age): Influence

Although there are several items where large DIF contrasts exist (>0.5 logits), the DIF plot seems to follow a unidirectional pattern for all of the age groups, with the difficulty level of the items increasing or decreasing in the same direction for all groups. This could suggest that the differences found in the data can be ascribed to actual age group differences, rather than DIF.

Item INF8 showed statistically significant DIF across the comparison of the youngest age group with all three of the other age groups. The data suggests that it was easier for the 15–20-year-olds to endorse this item than it was for the other age groups. None of these comparisons yielded DIF contrast measures above 0.5 logits, suggesting that the impact of the DIF might only be specific to the current sample. Item INF1 showed significant DIF between the 15–20-year-old and both the 31–40 and 41–50-year-old groups. In both cases it was more difficult for the younger group to endorse the item than for their older counterparts. The effect of the DIF was large (>0.5 logits) suggesting that this item will cause age-related DIF in other samples as well, and might not be an appropriate item for participants without working experience. Item INF9 yielded statistically significant probabilities on the Mantel-Haenszel test for both the 15–20-year-old compared to 21–30-years-old and compared to 41–50-year-old comparison groups. Neither of these comparisons resulted in DIF contrasts above the 0.5 logit estimate, suggesting that the impact of DIF on these items are not probable to repeat in other samples. Item INF4 was flagged as potentially causing DIF among participants from the 15–20 and 31–40-year-old age groups, yielding a statistically significant Mantel-Haenszel test result. The relatively small DIF contrast (<0.5 logits) indicates that the impact of the DIF might not be present in other samples than the current sample. For the comparison between the 15–20-year-old and the 41–50-year-old sample groups item INF5 was flagged for potential DIF. Based on the difference in item difficulty for the two groups (>0.5 logits) and the statistically significant Mantel-Haenszel test result, it is probable that this item could also cause age-related DIF in future samples. Item INF7 was also identified as potentially causing DIF between 1–20-year-old and 21–30-year-old participants, and the effect of this DIF was large (>0.5 logits) with a statistical significance at the $p<0.05$ level. The younger participants found it easier to endorse this item than the participants with a little more exposure to the working environment.

Three items were identified for potentially causing DIF across both age group comparisons with the 21–30-year-old participants as the reference group, namely Items INF1, INF6 and INF7. The 21–30-year-old participants found it more difficult to endorse items INF1 and INF7, while they found it more difficult to endorse item INF6. The statistically significant Mantel-Haenszel test results suggest that these items will likely also cause DIF between specific age groups in other samples. For the 21–30-year-old the results indicated that item INF4 might potentially be causing DIF. Upon further investigation we see that the DIF contrast between the two groups doesn't exceed the 0.5 logit estimate, suggesting that the differences in item difficulty for the two groups are sample specific, and the impact of the DIF would not be significant in other samples. Item INF5 was further also identified as potentially causing DIF between the 21–30-year-old and the 41–50-year-old groups. This item also didn't yield a large DIF contrast, although it resulted in a significant probability, which could be influenced by sample specific elements.

None of the items from the Influence scale were identified for causing DIF between the 41–50 and 51–60-year-old age groups.

Gender

The DIF plot for gender on the Influence scale is represented in Figure 12. From the plot it can be seen that the DIF plots for the two genders follow a unidirectional pattern, suggesting that the items of the scale won't cause DIF in other samples.

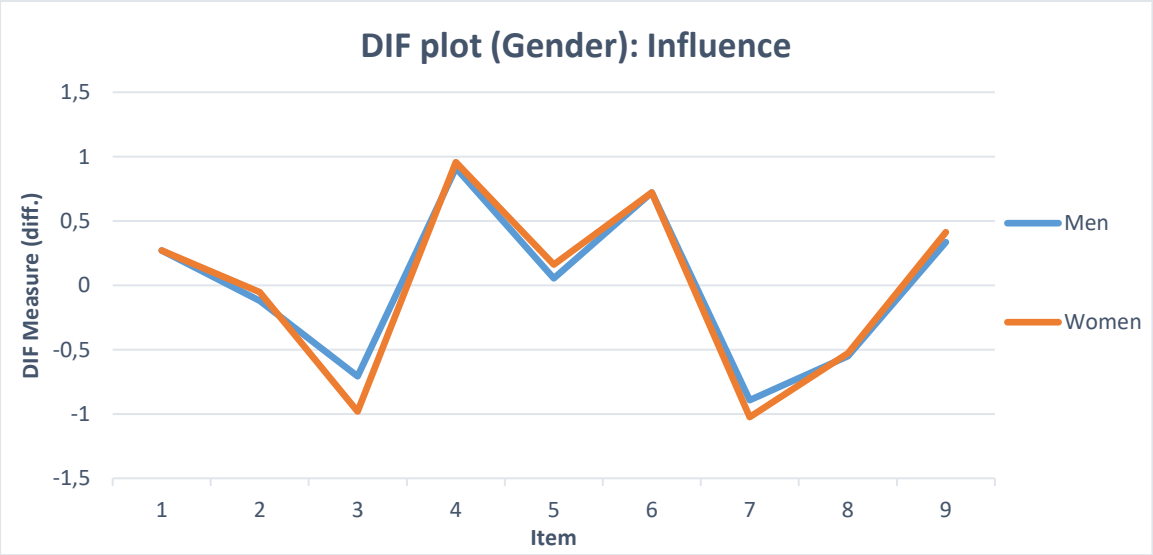

Figure 12. DIF plot (Gender): Influence

Only one item from the scale was identified for potentially causing DIF, yielding a statistically significant Mantel-Haenszel test result. The size effect of the DIF contrast however does not exceed the 0.5 logit estimate, suggesting that the DIF seen in the current sample will not likely be evident in other samples.

Ethnicity

The DIF plot for black and white participants (Figure 13) reveals a relatively unidirectional plot for both ethnic groups with only one item (INF7), resulting in a bidirectional plot point.

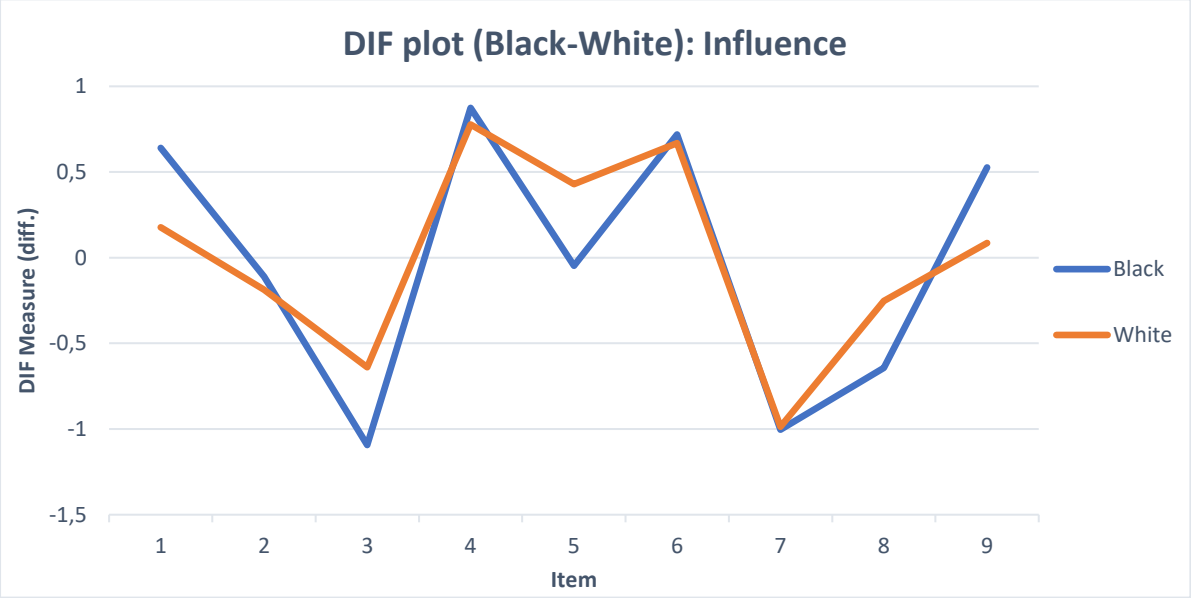

Figure 13. DIF plot (Ethnicity): Influence

Based on the DIF analyses 5 items were flagged for potential DIF between the two ethnic groups, but based on the small (<0.5 logits) DIF contrasts between the two groups, none of these items are likely to cause DIF due to racial difference, but is rather a result of the composition of the current sample.

Creativity

Item Fit

Based on the item difficulty/endorsability statistics, there is a fairly equal spread of item difficulty across the Creativity scale. None of the items were flagged for potential misfit based on their infit mean square statistics.

The item map for the Creativity scale (Figure 14) revealed that the majority of the sample found the items easy to endorse, and tended to rate the items as important, or very important to them in their careers.

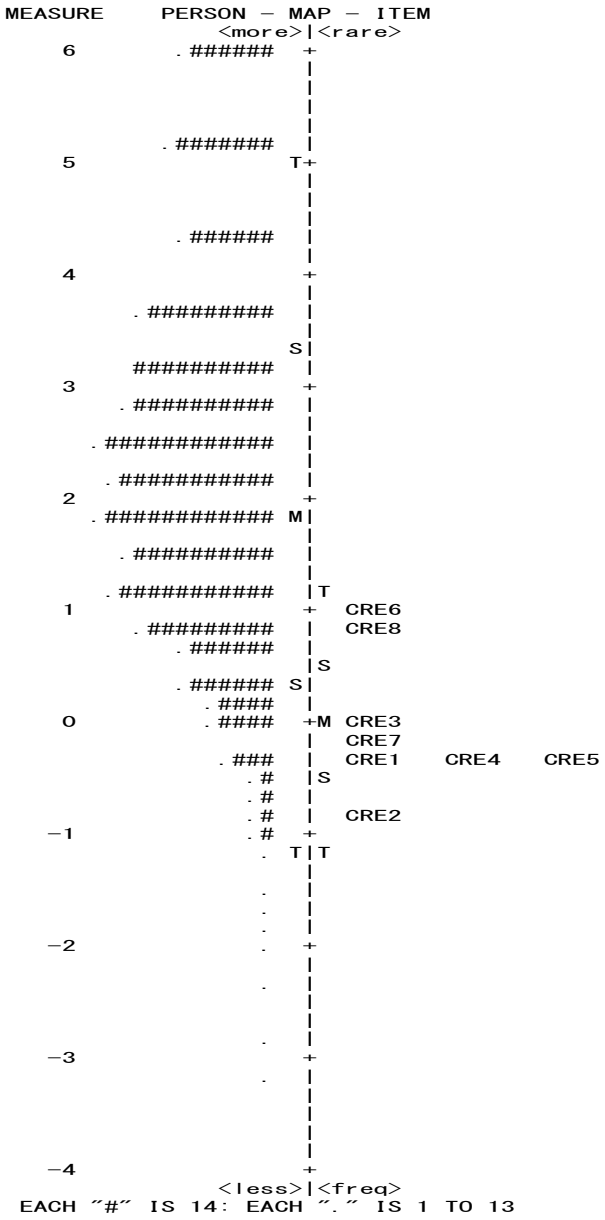

Figure 14. Item map: Creativity

Differential Item Functioning

Age

Figure 15 gives a graphical representation of the item difficulty for the various age groups on the Creativity scale. The plot reveals several items where large DIF contrasts exist between the various age groups.

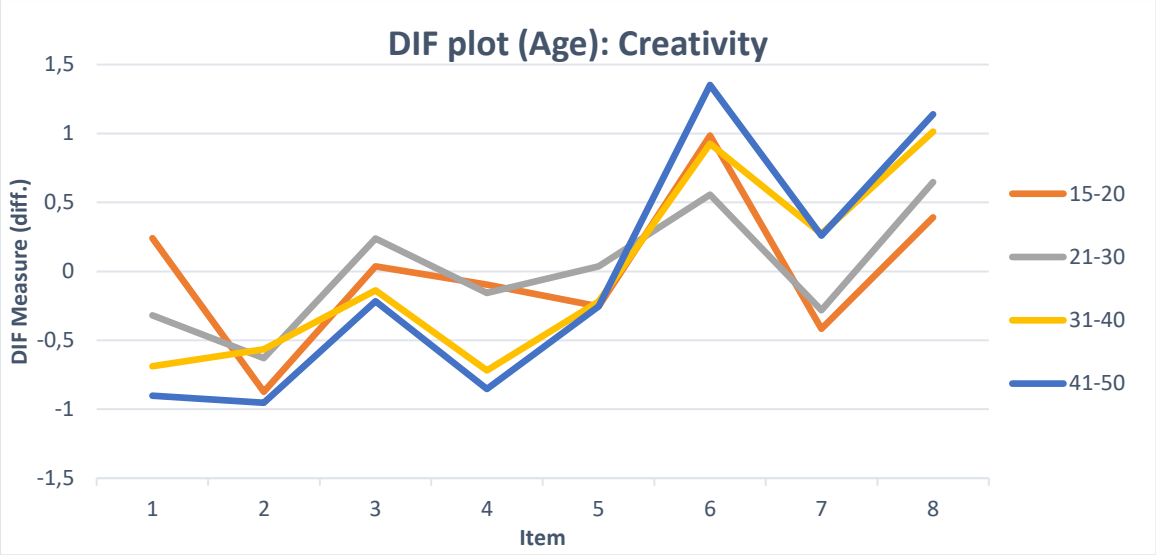

Figure 15. DIF plot (Age): Creativity

The results from the DIF analysis done with the 15–20-year-old participants as the reference group, and the other age groups as focal groups, showed that item CRE1 was flagged for potentially causing DIF across all three age group comparisons, with the youngest age group finding it more difficult to endorse that item than the older age groups. Across all three groups the DIF contrasts were statistically significant, suggesting that this item will likely cause DIF across different age groups in future samples as well.

For the comparison between the 15–20-year-old and the 21–30-year-old participants, two more items were flagged for potential DIF (items CRE5 and CRE6), but the DIF contrast for these items between the two groups did not exceed the 0.5 logit estimate, indicating that the impact of the DIF between these two age groups is not that extensive.

Items CRE4, CRE7 and CRE8 were identified as causing DIF (statistically significant at the p<0.5 level) between the 15–20-year-old participants, and both the 31–40-year-old and 41–50-year-old participants. The 15–20-year-olds found item CRE4 more difficult to endorse than the older participants, while items CRE7 and CRE8 seem to be easier for them to endorse. The large DIF contrasts and statistical significance, suggest that these differences are likely to also present in other samples, but it could be argued that the differences are due to actual generational differences and not due to the items themselves.

When comparing the 21–30-year-old group with the older age groups, items CRE4 and CRE7 showed statistically significant DIF across both age group comparisons, with the 21–30-year-olds finding item CRE4 more difficult to endorse than their older counterparts, and item CR7 easier to endorse. This is similar to the results from the 15–20-year-old groups, where we found that the younger population groups found the same items easier or more difficult to endorse. CRE3 was flagged as an item that can potentially cause DIF between 21–30-year-olds and 31–40-year-olds, but the DIF contrast between these groups didn't exceed the 0.5 logit estimate, indicating that the effect of the statistically significant DIF result is only relevant to the current sample, and the differences in item difficulty are a result of actual age differences and not item bias. Items CRE6 and CRE8 were identified in the comparison between 21–30-year-olds and 41–50-year-olds as potential DIF causing items. In both cases the analyses revealed statistically significant Mantel-Haenszel test results, but

only item CRE6 had a large DIF contrast. This suggests that item CRE6 will likely also cause DIF between age groups in other samples.

When looking at the DIF for the reference group 31–40-year-old participants and focal group – 41–50-year-old participants, only one item was flagged for causing DIF between the two age groups, but the impact of this DIF was small (<0.5 logits), suggesting that the probability of this item causing DIF in other samples is unlikely.

Gender

A graphical presentation of the DIF plot for men and women on the Creativity scale is presented in Figure 16.

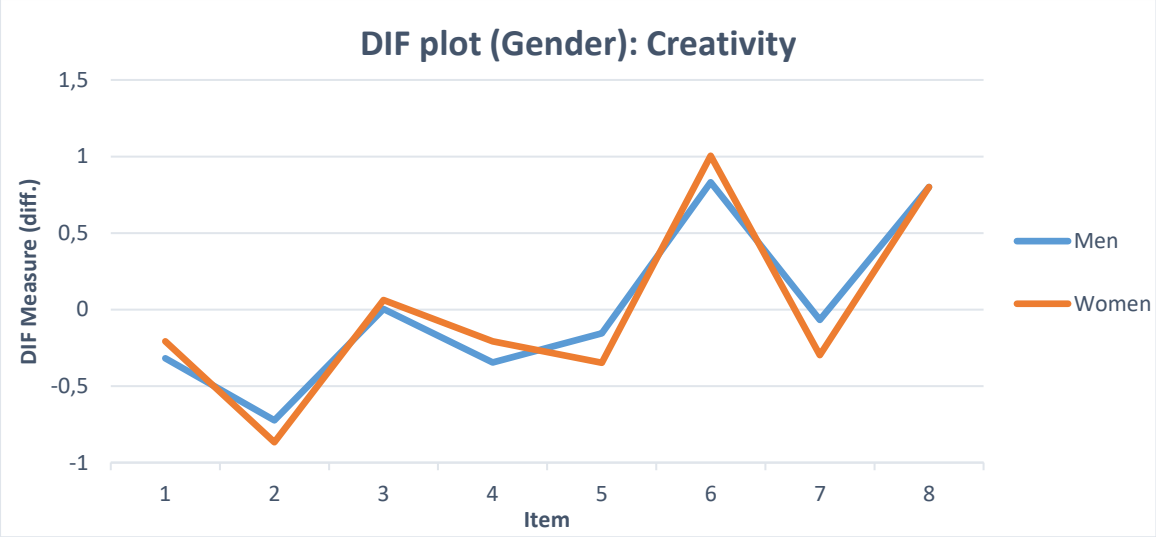

Figure 16. DIF plot (Gender): Creativity

The two DIF plots follow a seemingly unidirectional pattern with none of the items revealing large DIF contrasts between the two gender groups. This indicates that the items of the Creativity scale are not likely to cause gender-related DIF in other samples.

When investigating the DIF statistics for gender, five items were potentially causing DIF, but based on their small DIF contrasts, we concluded that the differences in response styles to these items are due to actual gender differences in the current sample and not due to item bias.

Ethnicity

The DIF plots for the two ethnic groups (Figure 17) indicate a mostly unidirectional pattern, with only one item showing a large DIF contrast.

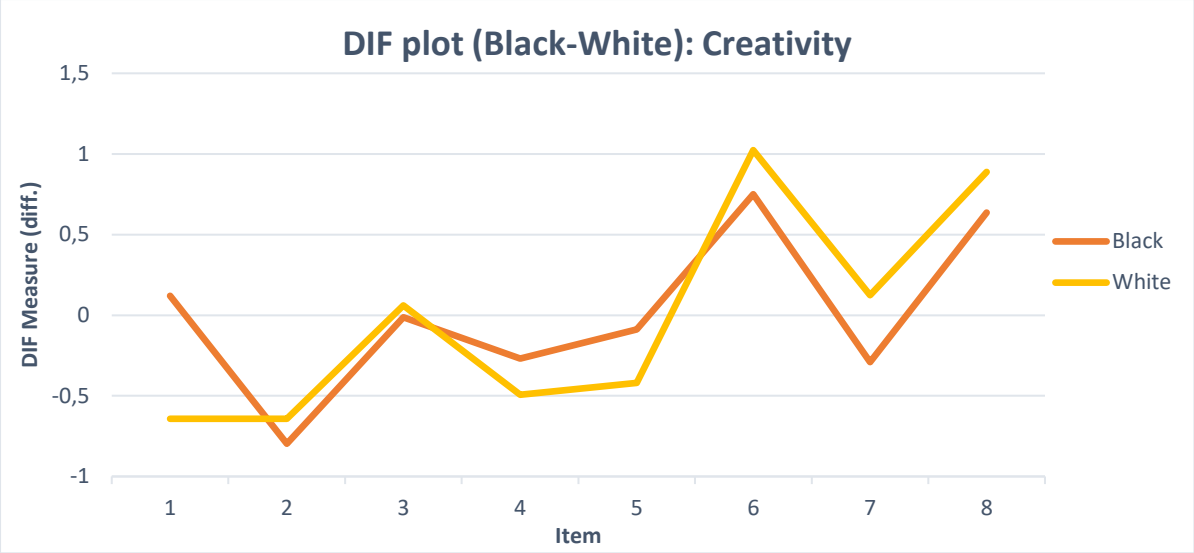

Figure 17. DIF plot (Ethnicity): Creativity

The statistics for the items flags item CRE1 as having a statistically significant DIF contrast larger than the 0.5 logit estimate. Item CRE8 is also flagged for yielding a statistically significant Mantel-Haenszel test result, but this item did not exceed the 0.5 logit estimate, indicating that the impact of the DIF contrast is not likely to repeat in other samples. Item CRE1 might, however, also cause Ethnicity DIF in other samples. We found that black participants found it more difficult to endorse this item than their white counterparts, and the item might need to be reviewed for future inclusion in the CVS.

Independence

Item Fit

Judging from the measure, statistics items from the Independence scale were less easy to endorse than those from the other CVS scales. We see a lot more items with positive fit statistics, suggesting that participants' ratings for the scale would be more equally spread between the 5 Likert-type options. None of the items from the scale were flagged for potential under- or overfit.

When looking at the item map for the Independence scale (Figure 18), there is a more equal spread between participants' ability to endorse the items and the endorsability of the items. There is however still a portion of participants who found all of the items from the scale too easy to endorse in a positive direction, i.e., ratings of 4 or higher.

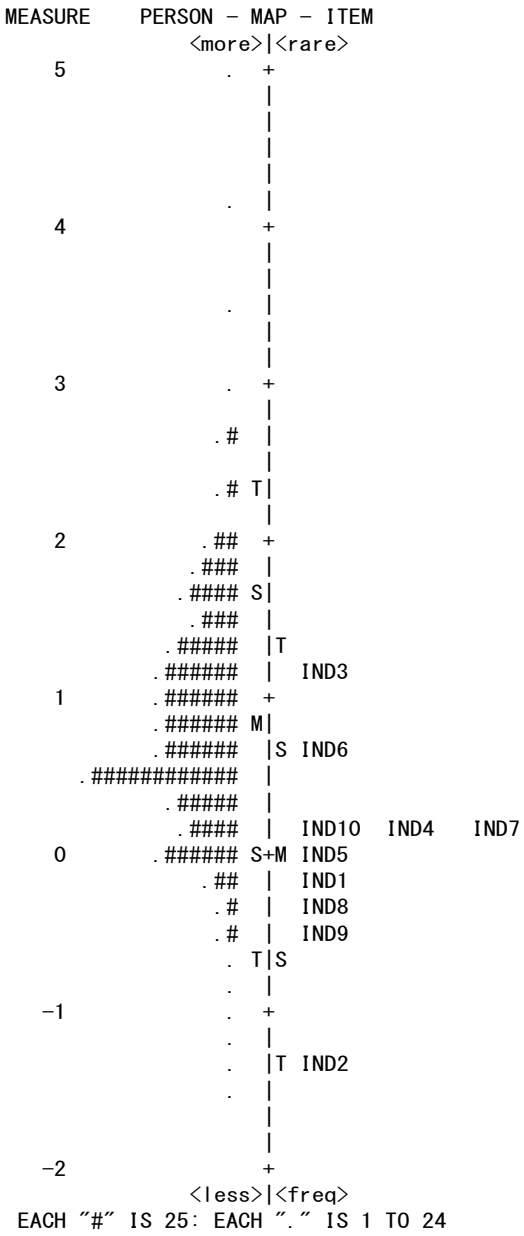

Figure 18. Item map: Independence

Differential Item Functioning

Age

The Independence scale age DIF plots are presented in Figure 19 and reveal several items with large DIF contrasts, and these will be discussed based on their statistical significance in the following sections.

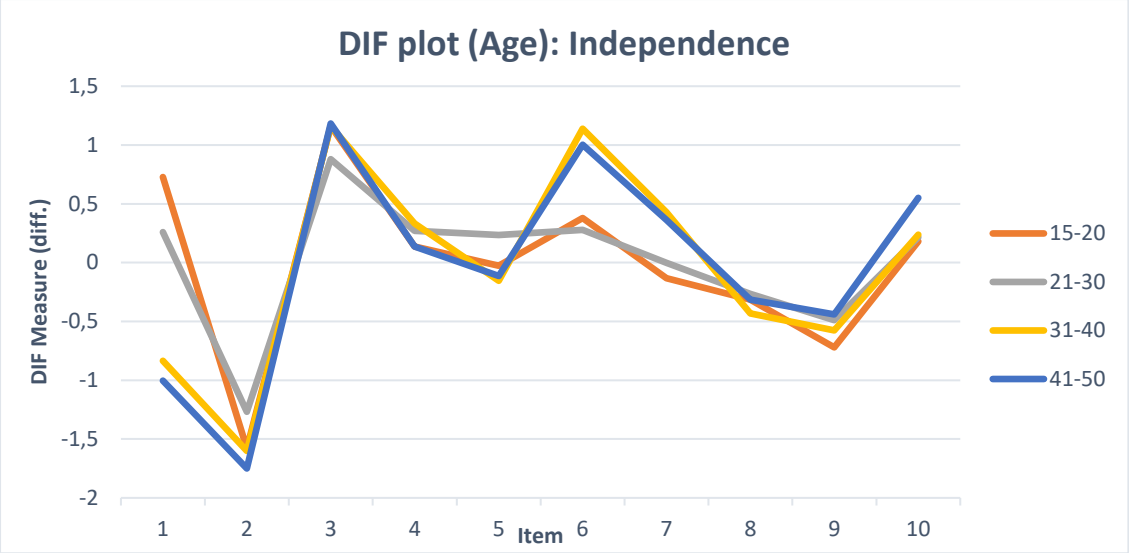

Figure 19. DIF plot (Age): Independence

For analyses where the 15–20-year-old group was used as the reference group with the other age groups as focal groups, we found that item IND1 was flagged for potentially causing DIF across all three the age group comparisons. In two of the comparisons (15–20-year-olds with both 31–40 and 41–50-year-olds) the DIF contrasts for this item was large (>0.5 logits) along with statistically significant DIF contrasts suggesting that this item will probably also cause age DIF in other samples. The DIF contrast in the comparison between the 15–20-year-old and 21–30-year-old participants was not large enough to consider this item as causing DIF in those age groups. Two more items were flagged for potential DIF for the 15–20 and 21–30-year-old age groups (items IND3 and IND5), but neither of these items yielded large enough DIF contrasts to be considered for DIF. Item IND5 also showed statistically significant results for the 15–20 and 31–40-year-old comparison, but again didn't have a large enough DIF contrast to be considered for DIF. Item IND6 and IND7 were flagged for potentially causing DIF in both the 15–20 with 31–40, and 15–20 with 41–50-year-old age group comparisons. In both groups the younger population group found it easier to endorse both items than their older counterparts.

We next provide an overview of the results from the DIF analyses run with the 21–30-year-old age group as the reference group. Based on the results items IND1, IND6 and IND7 were flagged for potentially causing DIF in both age group comparisons. The DIF contrasts for item IND7 were however small (<0.5 logits) and did not meet the criteria to consider the item for DIF. The 21–30-year-old participants found it easier to endorse item IND6, but found item IND1 more difficult to endorse than their older counterparts. This is in line with what we found when comparing the 15–20-year-old group to the older age groups as well, and it might suggest that the DIF seen in the results are actual age-related differences and not a result of item bias. Items IND3 and IND5 were also flagged for potential DIF in the comparison with the 21–30-year-old age group as the reference group and the 31–40-year-old participants as the focal group. Neither of these items managed to yield DIF contrasts above the 0.5 logit estimate, suggesting that the impact of the DIF is specific to the current sample and not likely to be present in other samples.

Lastly, we also compared the 31–40-year-old and 41–50-year-old participants and found that none of the items from the Independence scale were flagged for potentially causing DIF between these two age groups.

Gender

The gender differences in terms of DIF measures are plotted in Figure 20.

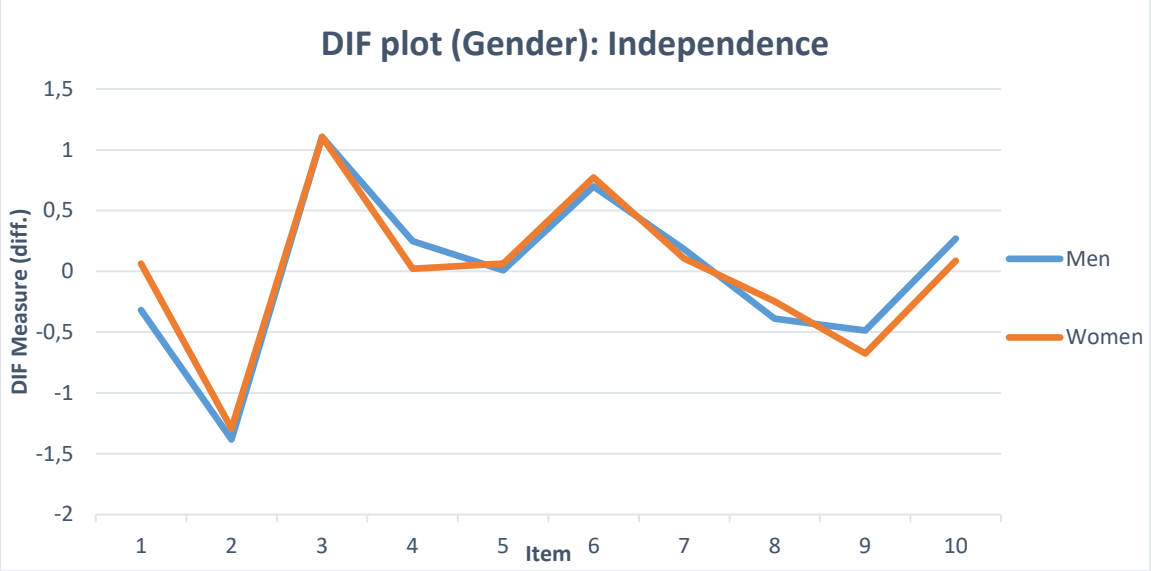

Figure 20. DIF plot (Gender): Independence

Based on the DIF plots, we see that the gender groups follow a unidirectional pattern, and based on the results from the DIF analysis, we see that although several items were flagged with statistically significant probabilities for causing DIF, none of them yielded large enough DIF contrast. It is therefore argued that none of the items in the Independence scale seem to be bias toward either gender groups.

Ethnicity

The DIF plot for black and white participants (Figure 21) indicates several items with large DIF contrasts that need to be investigated for potential ethnicity-related DIF. Overall, the two ethnic groups follow a unidirectional pattern, suggesting that the DIF contrasts might be related to actual differences between the groups and not due to item bias.

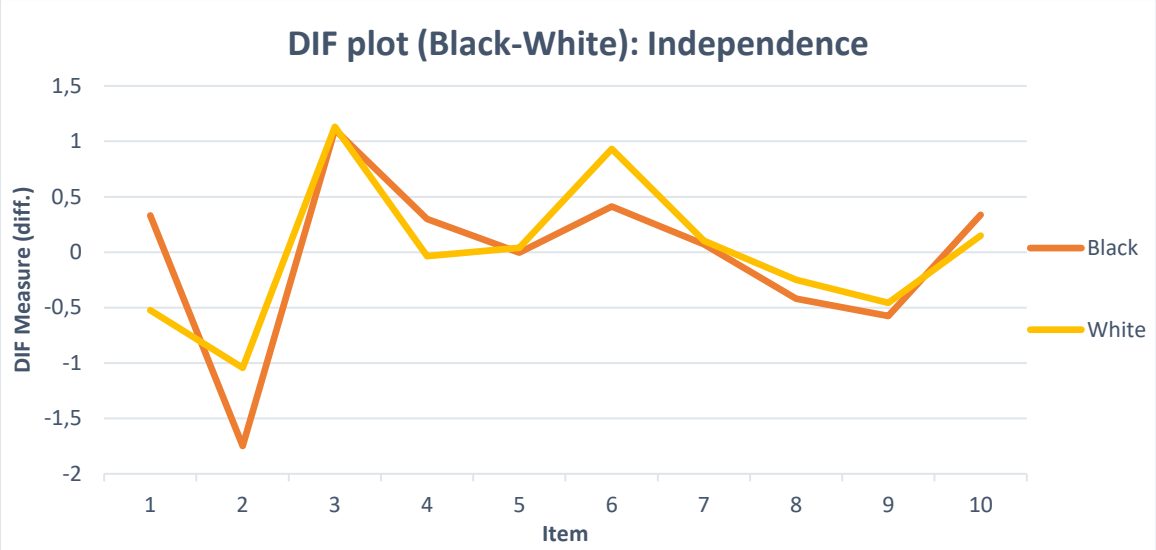

Figure 21. DIF plot (Black-White): Independence

In order to investigate these items, a DIF analysis was run between the two groups, with the black participants as the reference group. Four items were flagged for potential DIF, but item IND4 did not yield a large enough DIF contrast to be considered for DIF. The remaining three items – IND1, IND2 and IND 6 – yielded both large DIF contrasts and statistically significant Mantel-Haenszel test probabilities, suggesting that they are likely to cause ethnicity-related DIF in other samples too. When looking at the items it seems as if black participants found items IND2 and IND6 easier to endorse than their white counterparts, but found item IND1 more difficult to endorse. These items might need to be reviewed for further inclusion in the CVS, and based on the results from the ANOVA ran on ethnic groups where a statistically significant difference (with a small effect size) was found between black and white participants' scores on the Independence scale, it might require further investigation into the item-specific differences. It is also important to note that both items IND2 and IND6 failed to load with the other Independence items during the factor analysis, suggesting that these items might be biased to specific groups.

Excitement

Item Fit

The measure statistics suggest that there is an even spread between more difficult and easier to endorse items in the Excitement scale. Based on the results from the infit and outfit test, none of the items are flagged for potential under- or overfit.

The item map suggests that there was a more even spread between participants' tendency to endorse items positively, and the endorsability of the items for the Excitement scale. A large portion of the sample did still tend to rate items in this scale towards the higher ends of the Likert-type range spectrum.

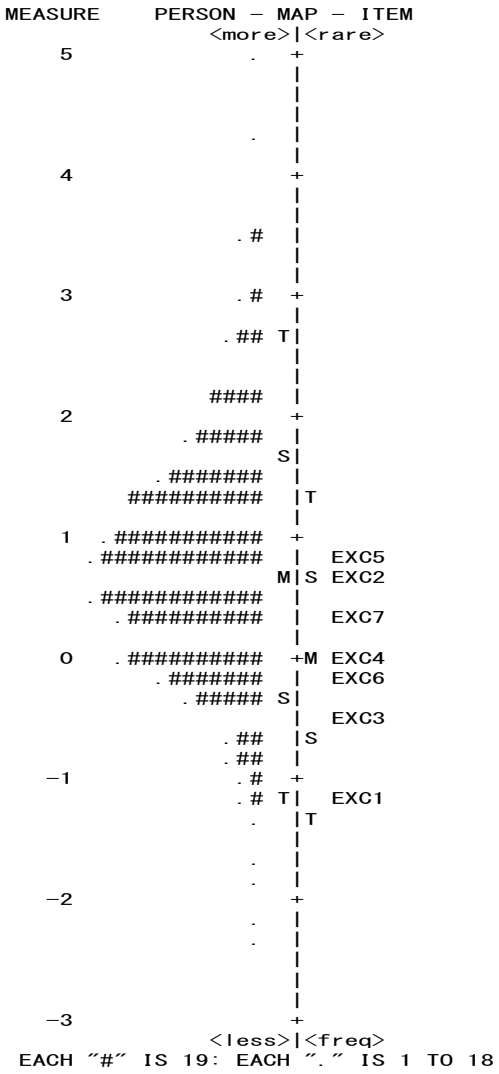

Figure 22. Item map: Excitement

Differential Item Functioning

The DIF plots for the different age groups are presented below. Figure 23 shows that there are several items in the Excitement scale that appears to have large DIF contrasts for the various age groups. In order to better understand this, we ran DIF analysis across the age groups using each age group as the reference group, with the other age groups as focal groups.

Age

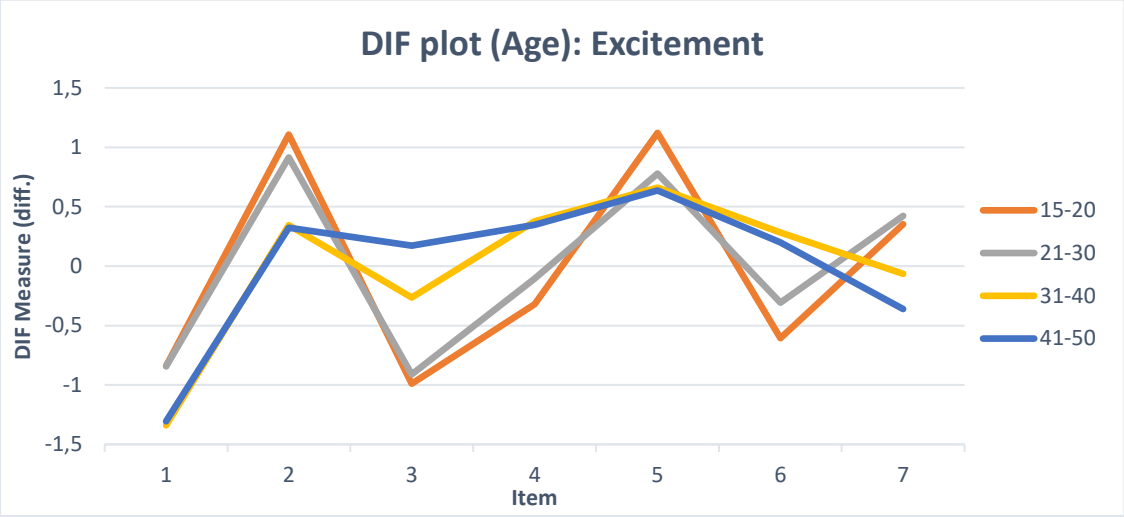

Figure 23. DIF plot (Age): Excitement

We first provide an overview of the results where the 15–20-year-old age group was used as the reference group. The results indicate that several items were flagged for potentially causing DIF. Although item EXC5 was flagged across all three of the age group comparisons, it didn't yield a large enough DIF contrast (>0.5) to be considered for DIF. For the comparisons between the 15–20-year-old participants and both the 31–40 and 41–50-year-old participants, all of the items from the Excitement scale were flagged for potential DIF. As already discussed, item EXC5 did not meet the requirements for DIF. Item EXC6 was also flagged for potential DIF in the 15–20-year-old and 21–30-year-old comparisons, but did not have a large DIF contrast. The younger participants found it easier to endorse items EXC3, EXC4 and EXC6, while they found items EXC1, EXC2, EXC5 and EXC7 more difficult to endorse. Items EXC1 and EXC7 also didn't load with other Excitement items during the factor analysis. Due to the fact that the entire scale is flagged for DIF between these population groups, it could also suggest that the scale is not appropriate for the younger population group (15–20-year-old participants), as they don't have sufficient exposure to environments that allow for these item-specific values.

Next, we provide an overview of the DIF results where the 21–30-year-old group was used as the reference group. Similar to the previous DIF table, most of the items were flagged for potentially causing DIF between the different age groups. Item EXC1 didn't yield a large enough DIF contrast in the second group comparison, while item EXC4 didn't have a sufficient DIF contrast in either groups, and item EXC7 failed to show a large enough DIF contrast in the first comparison. Of the remaining items that were flagged, the younger group found it more difficult to endorse item EXC2, but found items EXC3 and EXC6 easier to endorse. There is a statistically significant probability that these items will also cause DIF in other samples.

When looking at the DIF statistics where the 31–40-year-old age group was the reference group, two items were flagged for potential DIF, but neither of these items showed large DIF contrasts, suggesting that the impact of the DIF is specific to the current sample and not probable to repeat in other samples.

To note: The Excitement scale showed very low reliability for the younger age groups (<0.60) which could indicate that this scale is not appropriate for use with younger participants. The items in the scale also show little variation in the elements of the construct that it taps into, and with only 7

items, it can be expected that the scale won't perform as well as other scales in the CVS. This could have influenced the results from the DIF analysis, and it is very possible that similar results will not be obtained in samples that exclude participants of ages 15–17.

Gender

Figure 24 provides the graphical representation of the DIF plots for men and women on the Excitement scale. The plots follow a unidirectional pattern which suggests that there is not sufficient evidence for DIF in any of the current items. To further investigate the DIF contrast and probability for causing DIF, we ran a DIF analysis across the gender groups. Although seven of the items yielded statistically significant Mantel-Haenszel test results, none of them had sufficient item DIF contrasts (>0.5 logits) to substantiate DIF in other samples.

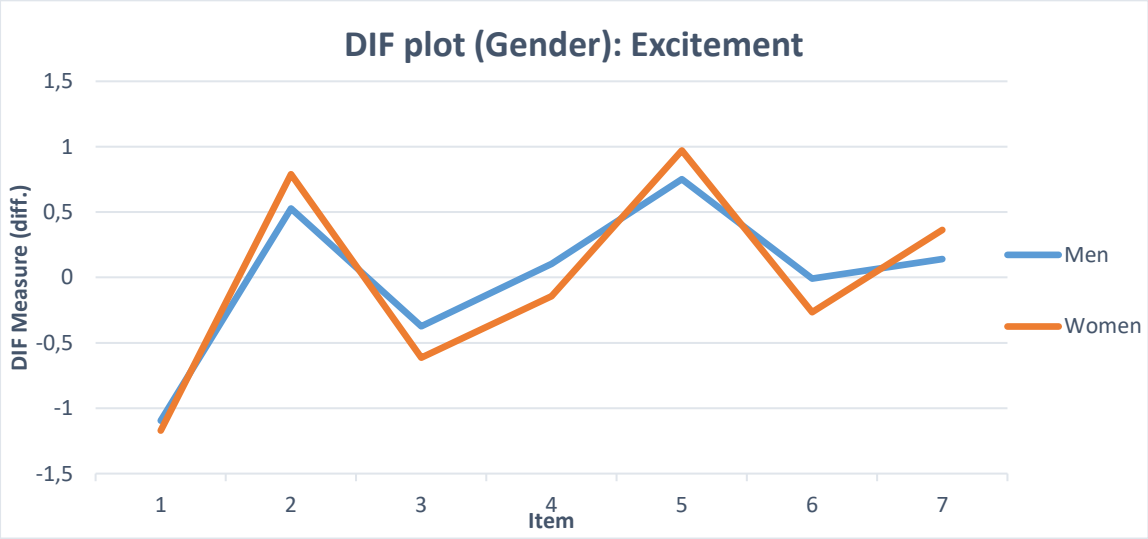

Figure 24. DIF plot (Gender): Excitement

Ethnicity

Figure 25 shows the two ethnic groups' DIF measures plotted for the Excitement scale. From the graph it is clear that several items had large DIF contrasts and might potentially cause ethnicity related DIF in other samples. The statistics from the DIF analysis ran between black and white participants, flagged all of the items for potential DIF based on their probability at the $p<0.05$ level, but only items EXC1, EXC3 and EXC4 yielded DIF contrasts large enough to consider DIF. Black participants found item EXC1 more difficult to endorse, but seemed to have found items EXC3 and EXC4 easier to endorse than the white participants. The statistically significant Mantel-Haenszel test results, along with the large DIF contrasts for these items, indicate that these items are likely to cause ethnicity-related DIF in other samples as well. However, due to the fact that the direction of DIF was not the same across the scale, it is unlikely that this will translate to the test level.

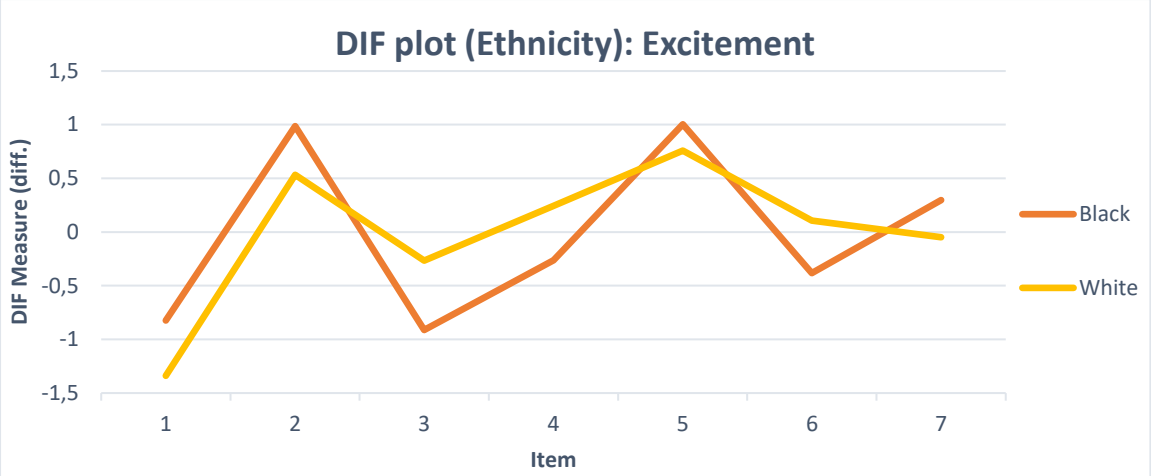

Figure 25. DIF plot (Ethnicity): Excitement

Career Development

Item Fit

There seems to be an equal spread of items that are easy and difficult to endorse, with one specific item being flagged as relatively difficult to endorse (CAD1). Based on the mean square statistics only 1 item is flagged for potential underfit, suggesting that item CAD1 might be measuring something different from the rest of the items in the scale. The item was further also flagged in the reliability analysis as having very low correlation with any of the other scale items (<0.10). It further also seemed to lower the reliability of the entire scale overall. It might be worthwhile considering the removal of this item.

This is also clear from the item map (Figure 26) where we see that most of the participants from the current sample find the items easy to endorse and will tend to lean towards higher ratings for most of the items in the scale.

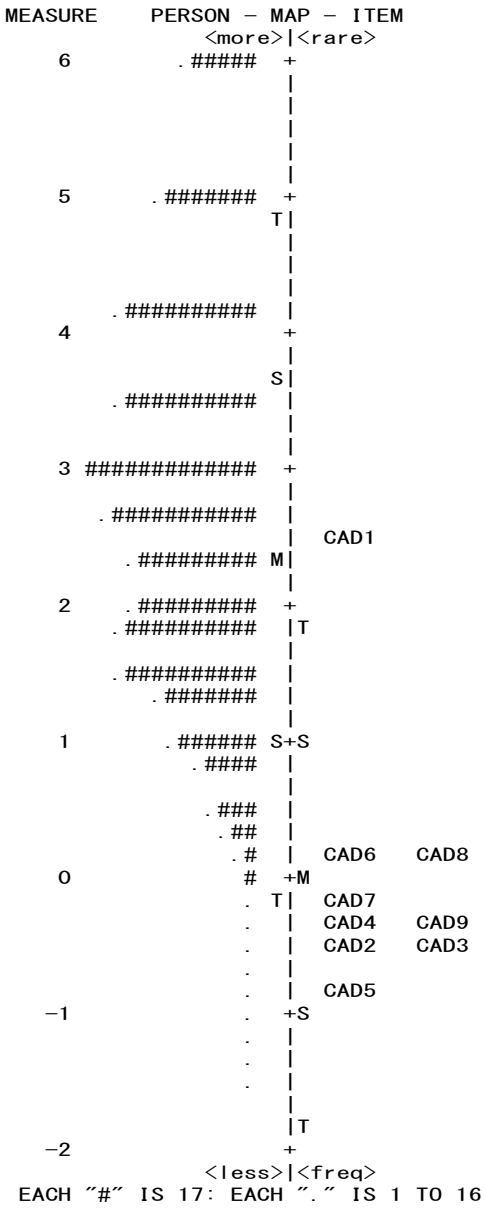

Figure 26. Item map: Career Development

Differential Item Functioning

Age

Figure 27 provides a graphical representation of the item difficulty for the various age groups on the Career Development scale.

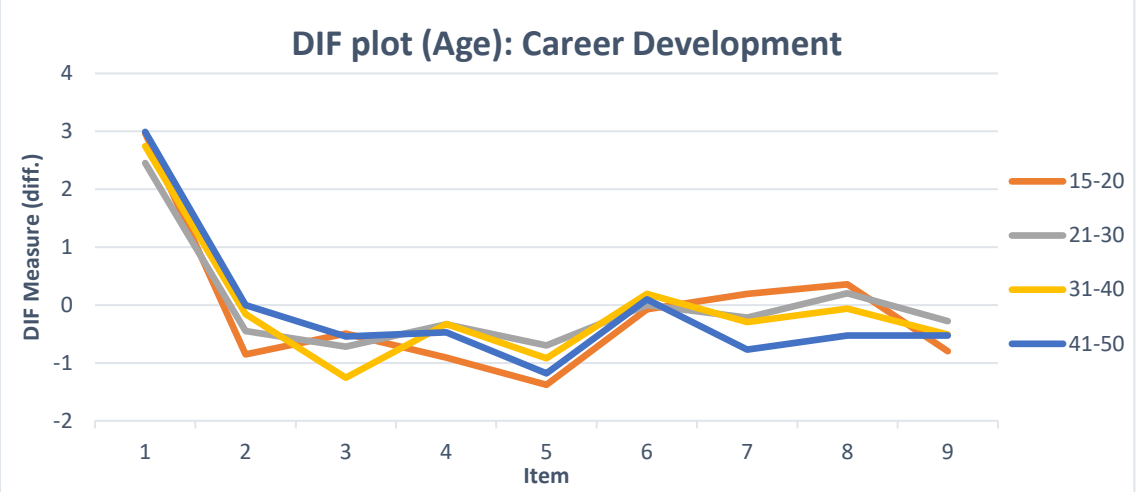

Figure 27. DIF plot (Age): Career Development

The DIF plot follows a relatively unidirectional pattern for all of the age groups, with only the 15–20-year-old groups deviating largely (>0.5 logits) from the pattern. We first analysed the age group DIF with the 15–20-year-old group as the reference group. Several items were flagged for potential DIF across these age group comparisons. For the comparison between the 15–20-year-old and 21–30-year-old participants, four items were flagged for potentially causing age-related DIF. The younger participants found it more difficult to endorse item CAD1, but easier to endorse items CAD4, CAD5 and CAD9. Based on the statistical significance of the DIF contrasts between these age groups, it is likely that these items will also cause age-related DIF in other samples. Item CAD4 was also flagged for causing significant DIF in the other two age group comparisons, but did not yield a large enough DIF contrast in the comparison between 15–20-year-old and 41–51-year-old participants. In the comparison between 15–20-year-old and 31–40-year-old participants, items CAD2, CAD3, CAD4 (as indicated already), CAD5, CAD7, and CAD8 were flagged for potential DIF, but only items CAD2, CAD3, and CAD4, yielded sufficient DIF contrasts. The younger age group found it more difficult to endorse item CAD3, and easier to endorse item CAD2 than their older counterparts. The 15–20-year-old participants found item CAD2 easier to endorse than the 41–50-year-old participants, and found items CAD7 and CAD8 more difficult to endorse. These three items showed large DIF contrasts, and significant Mantel-Haenszel test results indicating that they will probably also cause DIF in other samples where there are age differences.

Based on the results of the DIF analysis where the 21–30-year-old age group was the reference group, only two items were flagged for DIF and only between the 21–30-year-old and 41–51-year-old groups.

When looking at potential DIF between 31–40 and 41–50-year-old participants, only one item had a statistically significant Mantel-Haenszel test result, but it didn't show a large enough DIF contrast to be considered as also potentially causing DIF in other samples.

Gender

Figure 28 represents the item difficulty across the Career Development scale for men and women. The two plots follow a unidirectional pattern suggesting that the likelihood of item bias is minimal for gender differences.

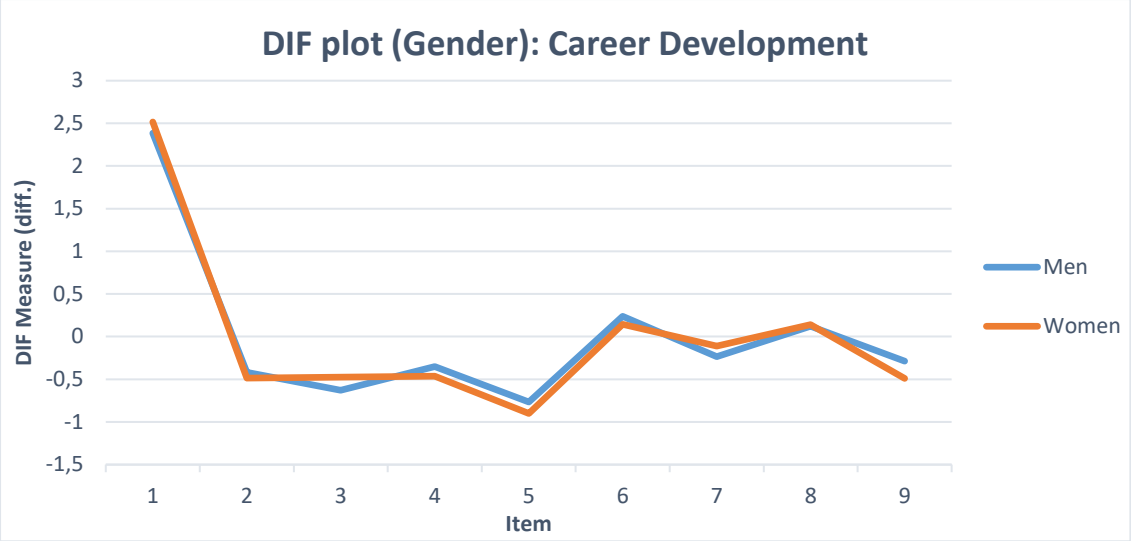

Figure 28. DIF plot (Gender): Career Development

When investigating the results from the gender DIF analysis, we found that although three items were flagged for potential DIF, their DIF contrasts weren't sufficient to suspect item bias, and that the significance is probably a result of the large sample size.

Ethnicity

We plotted the item difficulty for black and white participants in Figure 29 in order to give a graphical representation of the items with large DIF contrasts that might be causing DIF across ethnic groups.

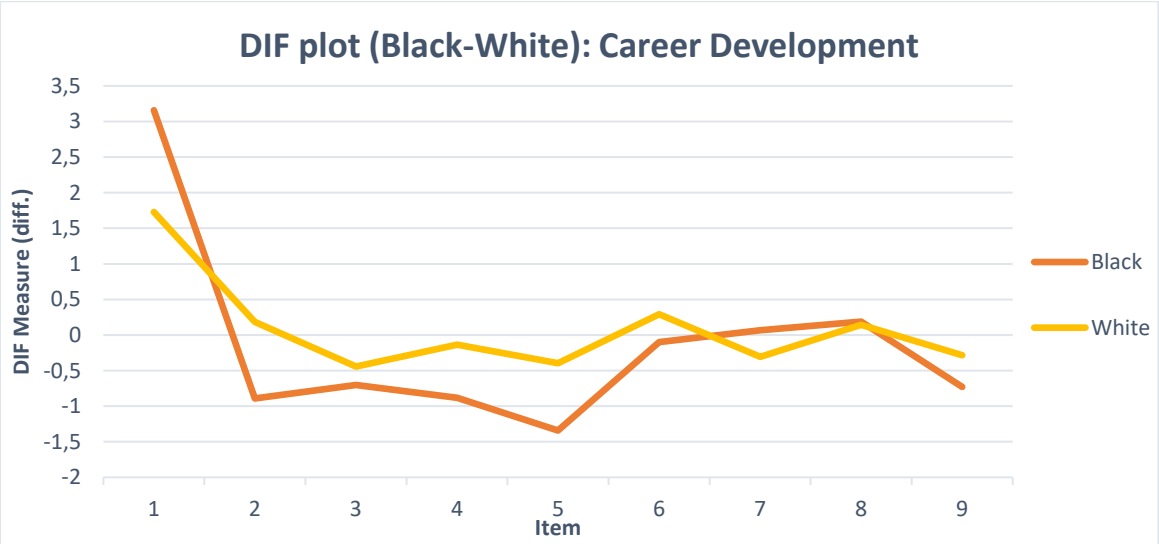

Figure 29. DIF plot (Ethnicity): Career Development

From the graph it can be seen that the plots didn't follow a unidirectional pattern, but rather deviated from this pattern for several of the items. In order to understand if these differences are due to DIF, we ran a DIF analysis, and the results showed that seven items were flagged for potential

DIF, but only four of these items (CAD1, CAD2, CAD4, and CAD5) had sufficient DIF contrast statistics to be considered for DIF. Black participants found it more difficult to endorse item CAD1, but easier to endorse items CAD2, CAD4, and CAD5.

Financial Rewards

Item Fit

When looking at the measure statistics we can see that the items are relatively equally spread between more difficult, and easier to endorse items. Item FIR6 was flagged as a potentially underfitting item, with a MNSQ value above the 1.35 acceptable range. This could suggest that this item is measuring a construct different to the rest of the items in the scale. During the factor analysis session, we found that this item didn't load with the other items from the scale, but instead loaded with items related to creativity and independence.

When investigating the item map for the Financial Rewards scale (Figure 30) where the participants' tendency to endorse an item is plotted against the endorsability of the items, it is clear that a large portion of the sample would tend to rate most of the items on the scale as either important, or very important, on the Likert-type rating scale.

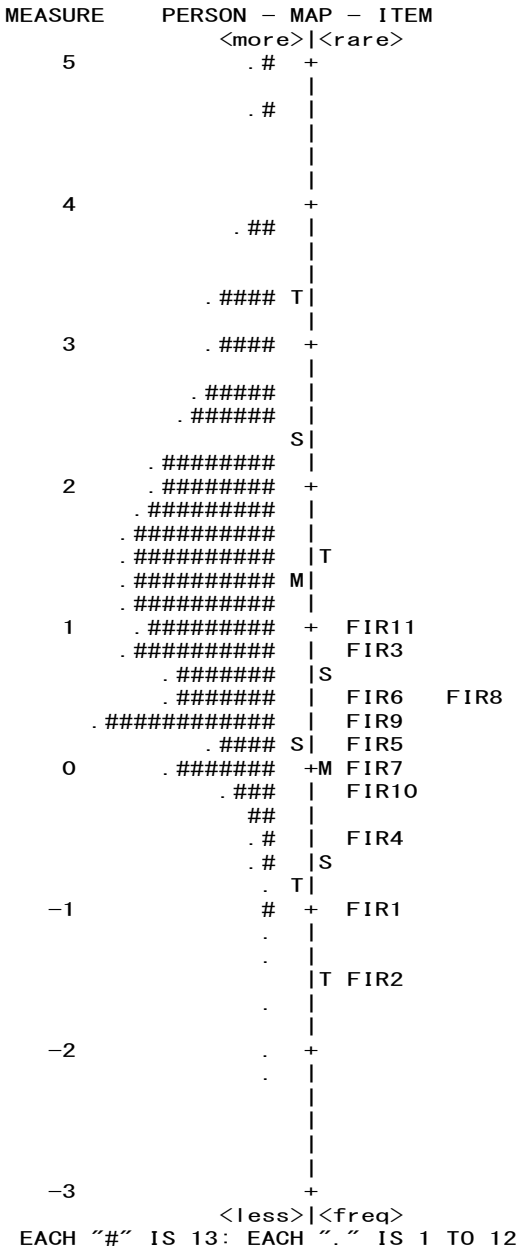

Figure 30. Item map: Financial Rewards

Differential Item Functioning

Age

The DIF plots for different age groups across the items of the Financial Rewards scale are shown in Figure 31. For the most part the plots follow a unidirectional pattern, with the two younger age groups deviating slightly from the two older age groups for specific items.

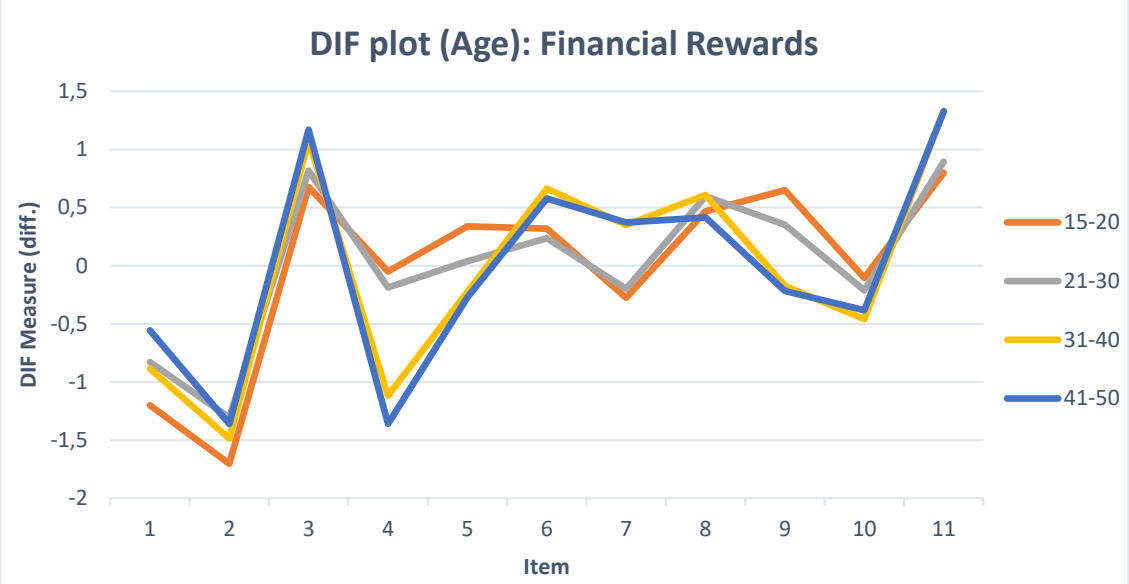

Figure 31. DIF plot (Age): Financial Rewards

Where the 15–20-year-old group was the reference group, only one item (FIR5) was highlighted for potentially causing DIF among 15–20 and 21–30-year-olds, but the DIF contrast was not large enough for the effect to be considered transferable to other samples. When looking at the 15–20 and 31–40-year-old groups' results from the DIF analyses, several items were flagged with statistically significant results, but only items FIR4, FIR5, FIR7, FIR9, and FIR11 had large enough DIF contrasts to be considered as resulting in significant DIF. The younger participants found it easier to endorse items FIR7 and FIR11, while their older counterparts found items FIR4, FIR5, and FIR9 easier to endorse. Items FIR1, FIR4, FIR5, FIR7, FIR9, and FIR11 all showed large DIF contrasts and statistically significant Mantel-Haenszel test results for the 15–20 and 41–50-year-old participants group. For most of these items (except FIR1) the results are expected, as they were similar to those between the 15–20-year-old and 31–40-year-old groups.

The findings of the DIF analysis with the 21–30-year-old age group as the reference group, showed that items FIR4, FIR7, FIR9, and FIR11 were flagged across both age group comparisons as yielding large DIF contrasts and statistically significant results.

The DIF analysis where the 31–40-year-old age group was the focus group, the results suggest that none of the items were flagged for potentially causing DIF among these two age groups.

Gender

The item difficulties for men and women on the Financial Rewards scale are plotted in Figure 32.

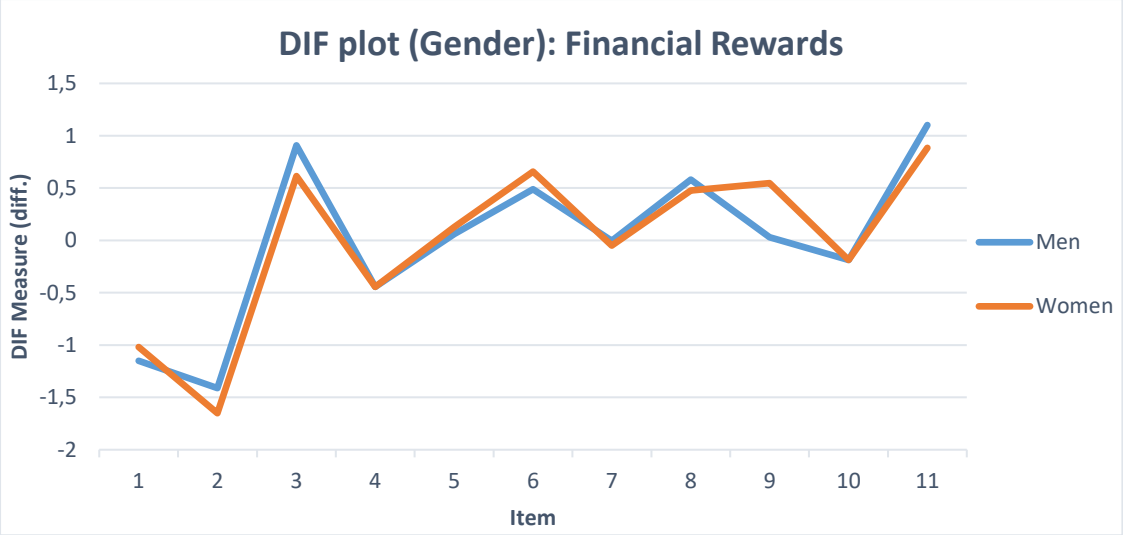

Figure 32. DIF plot (Gender): Financial Rewards

When inspecting the DIF plot, it seems that the item difficulty for the two gender groups follow a unidirectional pattern, with only item FIR9 performing differently for the two groups. When looking at the results from the DIF analysis, this is the only item that yielded a sufficient DIF contrast to be considered for potentially causing DIF across gender groups. It seems that men find item FIR9 easier to endorse than the women in the current sample. Four other items were flagged for DIF, but did not yield large enough DIF contrasts, suggesting that the significance of the small DIF contrasts obtained might be a result of the large sample size.

Ethnicity

The DIF plot for the black and white participants is presented in Figure 33 and follows a seemingly unidirectional pattern for most items. This suggests that most of the items from the Financial Rewards scale are unlikely to cause DIF in other samples.

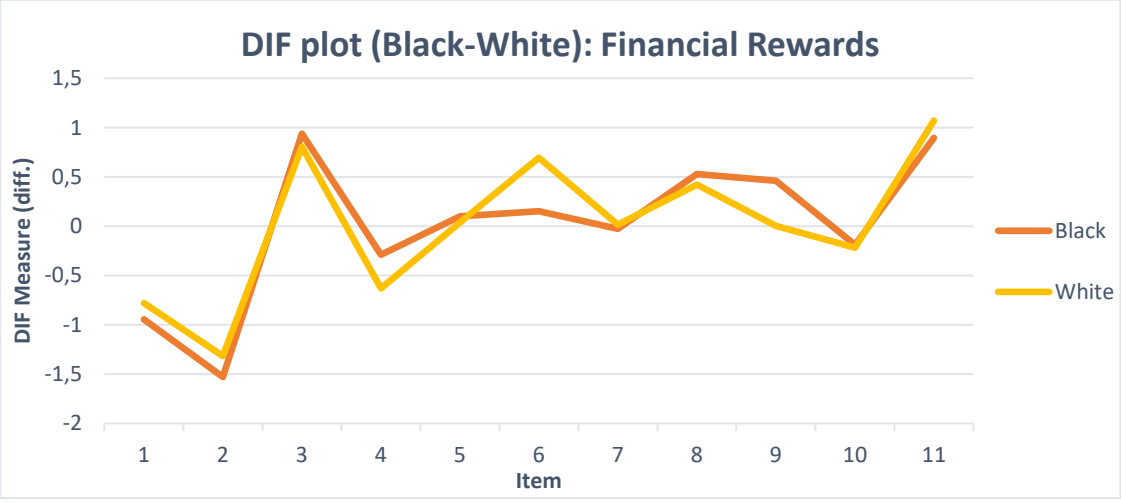

Figure 33. DIF plot (Black-White): Financial Rewards

Only one item showed a large enough DIF contrast paired with a statistically significant Mantel-Haenszel test result, namely item FIR6. Black participants found this item easier to endorse than their white counterparts.

Prestige

Item Fit

Based on the measure statistics it seems as if most of the items were relatively easy to endorse. None of the items from the Prestige scale were flagged for potential misfit based on their mean square statistics.

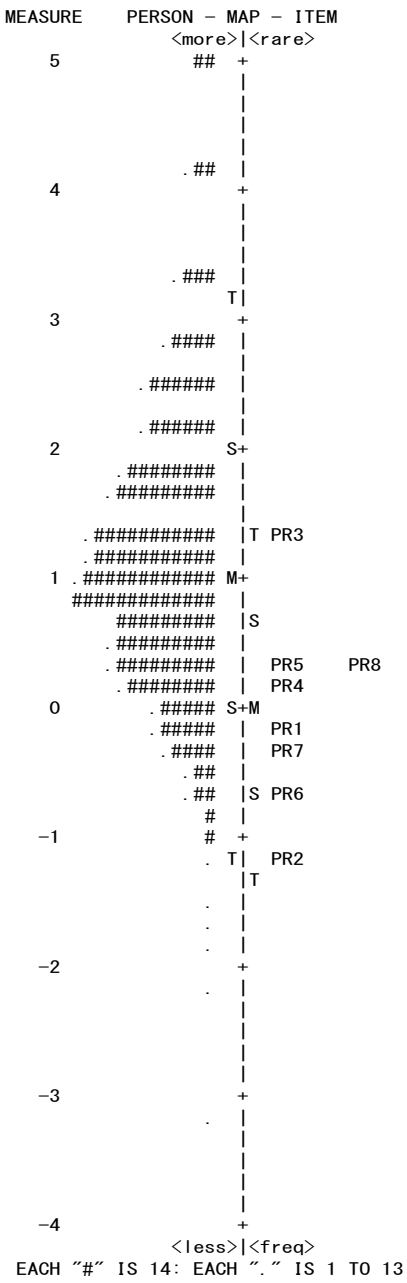

Figure 34. Item map: Prestige

The item map (Figure 9) suggests that the majority of participants from the current sample found most of the items (excluding item PR3) easy to endorse, and tended to rate the items a 4 or higher on the 5-point Likert-type scale.

Differential Item Functioning

Age

In order to investigate potential DIF among age groups on the Prestige scale, DIF analyses were done comparing each age group (reference group) with the other age groups in the sample. The overall item difficulty plots for each age group is presented in Figure 35, and it shows that the different plots follow a fairly unidirectional pattern with the larger DIF contrasts occurring between the youngest and oldest age groups.

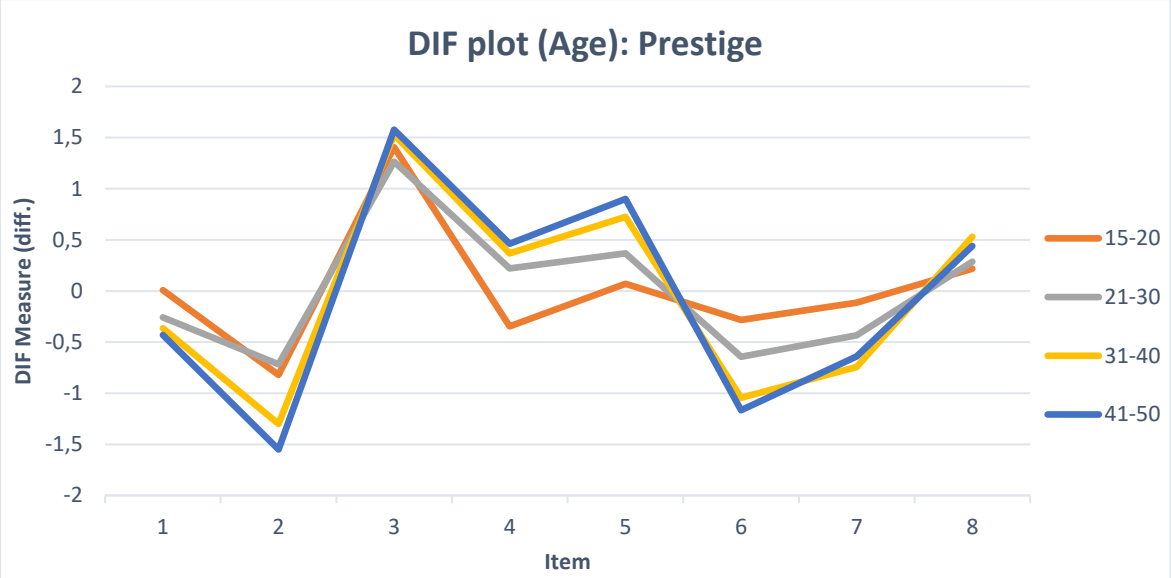

Figure 35. DIF plot (Age): Prestige

From the DIF analyses done where the 15–20-year-old age group was used as the reference group, it was seen that several items were flagged for potential DIF, but we only discuss those that yielded both statistically significant results, and large DIF contrasts (>0.50). For the comparison between the 15–20-year-old and 21–30-year-old participants, only item PR4 is considered for DIF. The younger participants found this item easier to endorse than their slightly older counterparts. This item was also flagged in the comparisons between the 15–20-year-old and other age groups with the same pattern, where the younger group found it easier to endorse the item. For the comparison between 15–20-year-old and 31–40-year-old participants, three more items were flagged for potential DIF: items PR5, PR6, and PR7. The younger participants found item PR5 easier to endorse, while older participants found items PR6 and PR7 easier to endorse. The same pattern is found in the comparison between 15–20-year-old and 41–50-year-old participants, with items PR5, PR6, and PR7 again being flagged for potential DIF. One other item, PR2, was also flagged between these specific age groups.

The findings from the DIF analysis with the 21–30-year-old group as the reference group, showed that item PR2 was flagged for DIF in both age group comparisons, with the older participants finding it easier to endorse the item than younger participants. This is similar to what we saw between the 15–20-year-old group and the older groups as well, suggesting that there might be age-related

differences in how participants choose to respond to this item. It could suggest real age-related differences, rather than item bias. Item PR5 was also flagged as potentially causing DIF between 21–30-year-old and 41–50-year-old participants in the current sample.

The DIF analyses of the final age group comparison showed that none of the items were identified for potential DIF between 31–40 and 41–50-year-old participants.

Gender

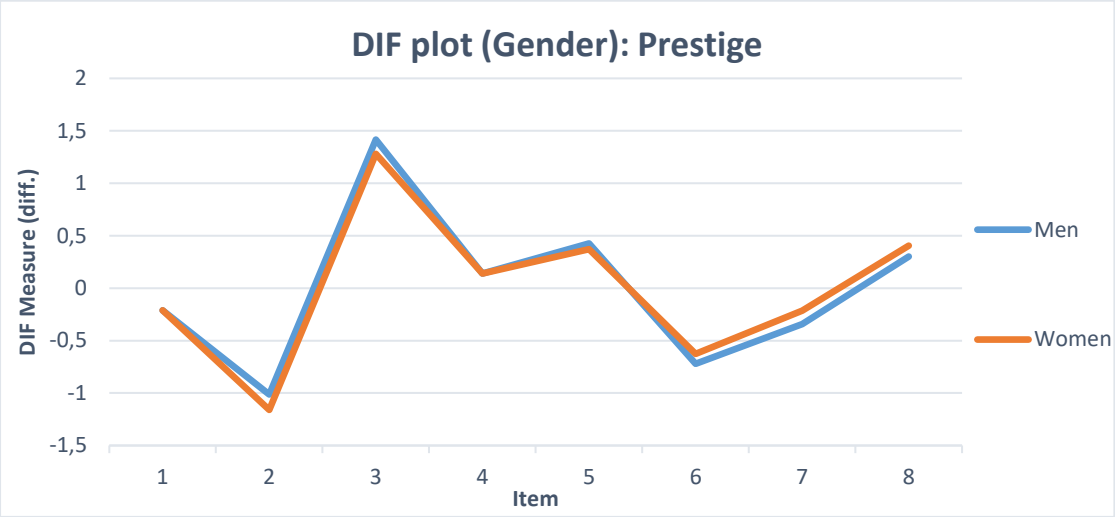

Figure 36. DIF plot (Gender): Prestige

Figure 36 and provides the results from the analyses conducted on the two gender groups for potential DIF across the Prestige scale items. From the graph it can be seen that the item difficulty plots for both genders follow a unidirectional pattern and the statistics revealed that none of the items yielded large enough DIF contrasts to be considered for potentially causing DIF.

Ethnicity

The Prestige ethnicity DIF plot is presented in Figure 37 and shows a unidirectional pattern for the black and white participants, suggesting that the items of the scale are not likely to cause ethnicity-related DIF between these groups in other samples.

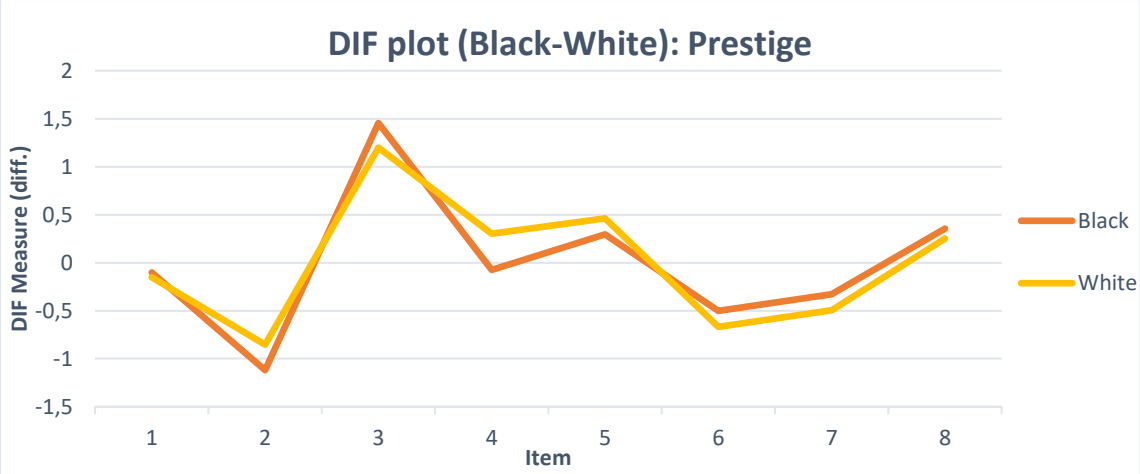

Figure 37. DIF plot (Ethnicity): Prestige

The results from the DIF analysis shows that none of the items were flagged for potential DIF across white and black participants.

Security

Item Fit

Based on the measure statistics, most of the items in the Security scale appear to be easy to endorse, i.e. participants will tend to favour the 'important' and 'very important' responses for most of the items. None of the items from the Security scale were flagged for potential misfit based on their mean square statistics.

When inspecting Figure 38, we see that the item map for this scale suggests that most of the participants found the majority of the items very easy to endorse. Only item SEC2 and SEC7 were more difficult to endorse, but a large portion of the current sample still found these items very easy to endorse. In other words, participants would tend to rate most of the items as being 'important' or 'very important' to them.

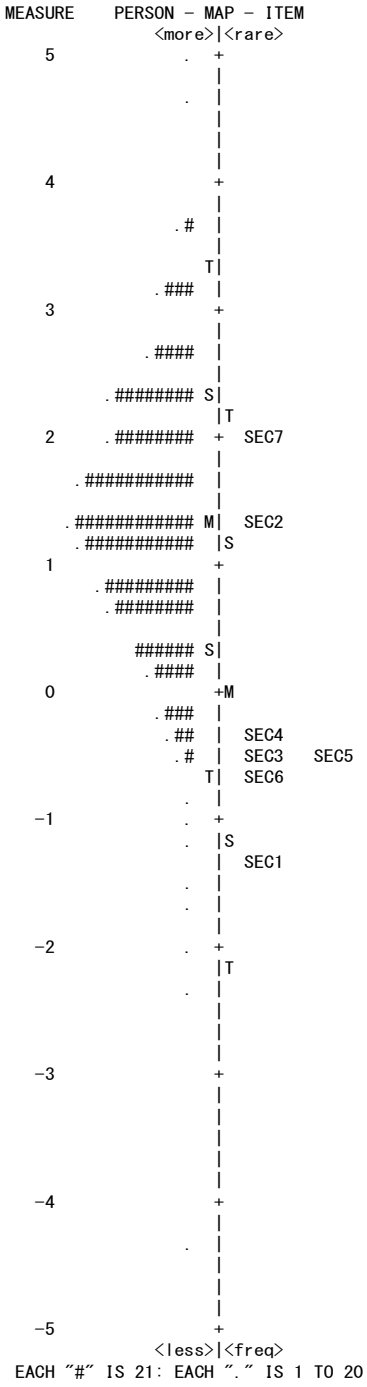

Figure 38. Item map: Security

Differential Item Functioning

Age

The item difficulties on the Security scale for the different age groups are plotted on the DIF plot presented in Figure 39. The DIF plots follow a fairly uniform pattern, with only the youngest age group deviating from the pattern for specific items.

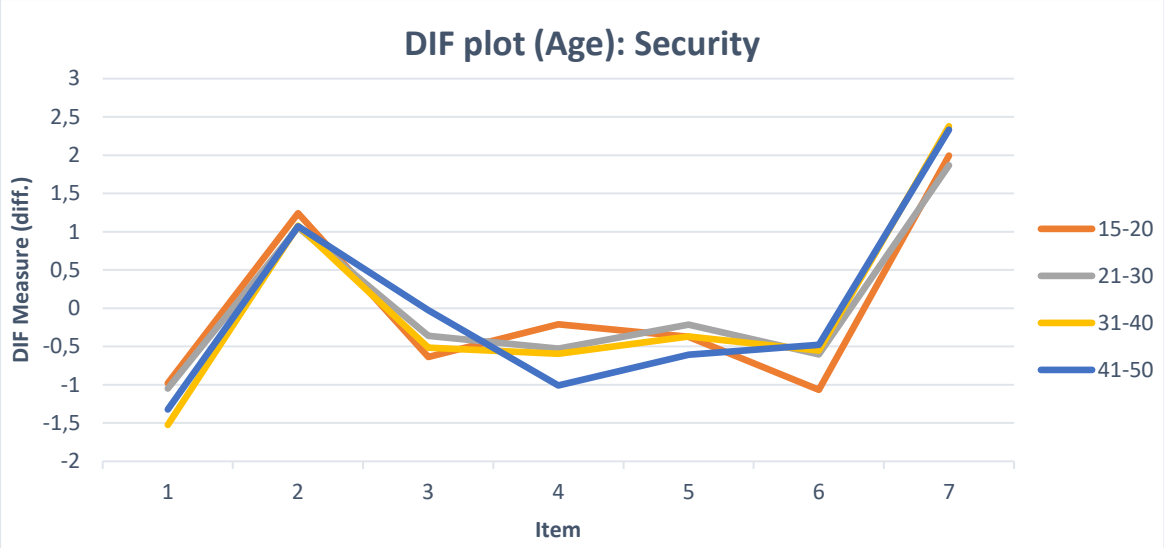

Figure 39. DIF plot (Age): Security

The DIF analysis with the 15–20-year-old age group as the reference group showed that although several items were flagged for potential DIF across the three age group comparisons, most of them didn't yield sufficient DIF contrasts to be considered for DIF in other samples. The only items that had both large DIF contrasts and statistically significant Mantel-Haenszel test results, were items SEC1 and SEC6 for the 15–20 and 31–40-year-old group comparison, and item SEC4 for the 15–20 and 41–50-year-old group comparison.

The 15–20-year-old participants found it easier to endorse item SEC6 than their older counterparts, but the older participants found it easier to endorse items SEC1 and SEC4. What is interesting to note is that item SEC1 loaded with items relating to Financial Rewards during the factor analysis phase, while items SEC4 and SEC6 also failed to load with the other Security items. This was also the scale with the lowest reliability for the 15–20-year-old sample, suggesting that perhaps these items are not suitable to the younger population group, and is likely to cause DIF between them and other age groups in future samples as well.

The results from the DIF analysis with the 21–30-year-old group as the reference group, highlighted only one item that might cause age-related DIF. The 21–30-year-old participants found it easier to endorse item SEC7 than the 31–40-year-old participants.

When looking at the final age group comparison (with the 31–40-year-old participants as the reference group), we found that although two items were flagged for potential DIF, neither of them yielded DIF contrasts large enough to be considered for DIF in other samples.

Gender

The gender DIF plot for the Security scale reveals a relatively unidirectional pattern, with only two items where there seem to be bigger DIF contrasts (Figure 40).

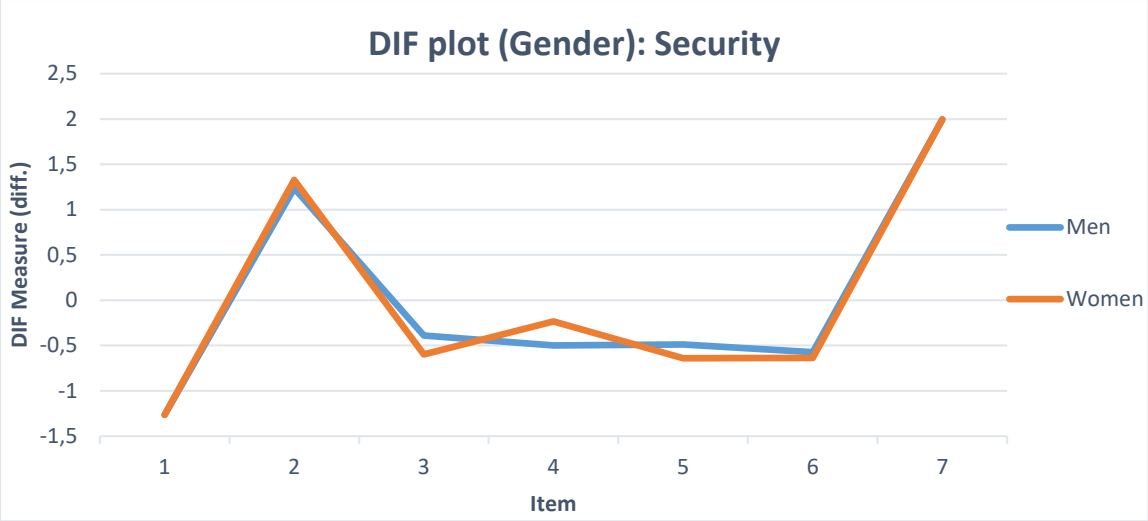

Figure 40. DIF plot (Gender): Security

When looking at the results from the DIF analysis, three items were flagged for potential DIF across gender groups, but none of these items had sufficient DIF contrasts to indicate that they will cause DIF in other samples. The statistically significant findings are most probably a result of the current sample size.

Ethnicity

The item difficulty for black and white participants is plotted in Figure 41 and reveals two items where the DIF plots do not follow a unidirectional pattern, namely items SEC5 and SEC6.

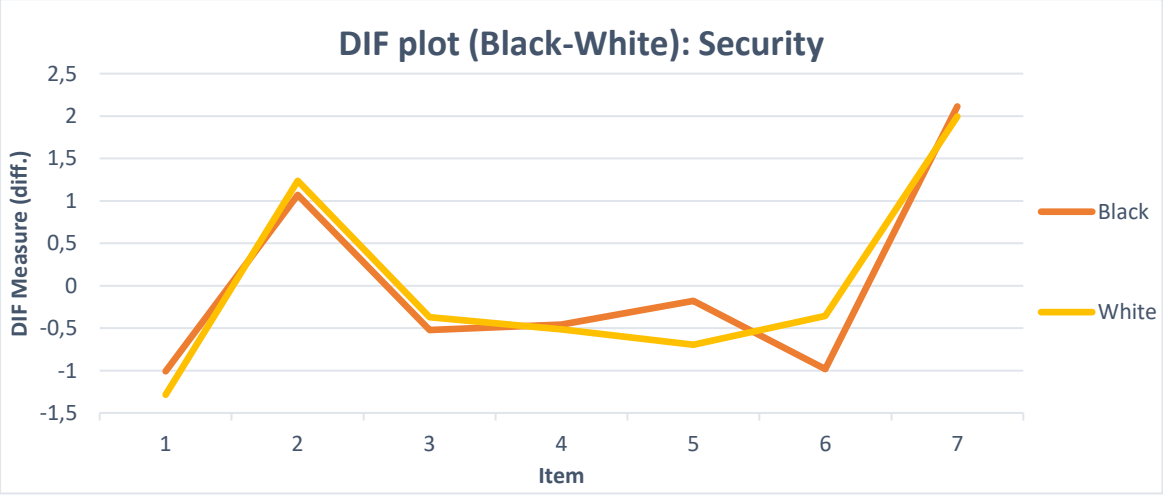

Figure 41. DIF plot (Ethnicity): Security

These items are flagged for potential DIF, based on the DIF plot, their large DIF contrasts (>0.50) and their statistically significant Mantel-Haenszel test statistics. Black participants found it easier to endorse item SEC6, while white participants found item SEC5 easier to endorse. It is likely that these items will also cause DIF between black and white participants in other samples.

Summary

New research about statistical significance proposed that the statistically significant *P*-value threshold should be adjusted from 0.05 to 0.005 (Benjamin, et al., 2017) and that this should only be one of the ways in which statistical significance is determined. If we applied the more stringent *p*-value and used the DIF contrast values as an estimate of effect size in order to highlight only those items that are likely to cause DIF in other samples, a summary of these items will look like Table 23.

In summary, we saw multiple items flagged for DIF across different age groups, with fewer items flagged for ethnic differences and gender differences. Although there were several highlighted items, the direction of the DIF did not seem to influence specific groups in the same direction overall. It can be argued that many of the differences found between age groups and ethnic groups, are due to actual differences between individuals, and not a function of the test items.

We ran the group comparison analysis across age groups without the 15-17-year-old participants, as the assessment is mostly intended for use with participants older than 18-years of age. We still found significant group differences across the age groups, as well as gender and ethnicity groups. This further suggests that there might be actual differences between groups that is accounted for by different sample groups placing more value on different work values. Specifically, with age-related differences, the level of exposure to the work environment, along the level of employment, could greatly influence what the individual will value in terms of their working environment. Although these differences could exist, we still argue that they should be in comparison to an overall sample group, and not just age-specific reference groups where some of the information about their values could be lost. With regard to ethnic differences, we argue that these may be a result of the historical background of South Africa, and that these differences actually reflect real cultural differences, rather than bias.

Table 23. DIF summary table

	Service Orientation	Team Orientation	Influence	Creativity	Independence	Excitement	Career Development	Financial Rewards	Prestige	Security
Gender										
Men and Women	SO5	-	-	-	-	-	-	FIR9	-	-
Ethnicity										
Black and White	-	-	-	-	IND1, IND2, IND6	EXC1, EXC3, EXC4	CAD1, CAD2, CAD4, CAD5	FIR6	-	SEC5, SEC6
Age										
15 - 20 & 21 - 30	-	-	-	-	-	-	CAD4, CAD5	-	PR4	-
15 - 20 & 31 - 40	SO3, SO6	TO1, TO3 TO9	INF1	CRE1, CRE4, CRE8	IND1, IND6, IND7	EXC2, EXC3, EXC4, EXC6	CAD2, CAD4	FIR4, FIR5, FIR7, FIR9, FIR11	PR4, PR5, PR6, PR7	-
15 - 20 & 41 - 50	SO2, SO6	TO1, TO2, TO3, TO9	INF1, INF8	CRE1, CRE4, CRE8	IND1, IND6	EXC2, EXC3, EXC4, EXC6, EXC7	CAD2, CAD7, CAD8	FIR4, FIR5, FIR7, FIR9, FIR11	PR4, PR5, PR6	SEC4
21 - 30 & 31 - 40	-	TO3	INF1, INF6, INF7	CRE4, CRE7,	IND1, IND6	EXC2, EXC3, EXC6	-	FIR4, FIR7	PR2	SEC7
21 - 30 & 41 - 50	-	TO1, TO2, TO3, TO9	INF1	CRE4, CRE6	IND1, IND6	EXC2, EXC3, EXC7	CAD8	FIR4, FIR7, FIR9	PR2, PR5	-
31 - 40 & 41 - 50	-	-	-	-	-	-	-	-	-	-

CONCLUSION

Overall, the CVS analyses yielded good psychometric properties, with acceptable reliabilities in the overall sample. Although there were some lower reliabilities in specific sample subgroups, the CVS is not used for selection purposes, but rather for guidance and development. Inferences made from the results of the CVS are not used for decision-making processes, and therefore the lower reliabilities are not a concern.

We saw several group differences between ethnic and age groups, and argue that these differences are due to actual group differences, rather than a function of test bias. When investigating the items (Rasch) we saw that the item endorsabilities across all of the scales were relatively similar for the different sample subgroups, and that the direction of the DIF were not uniform across the scales. We are also cognizant of the differences in sample sizes across the different subgroups, and argue that a general population norm would be sufficient as an itinerant norm until more data could be collected for further group comparisons.

REFERENCES

Benjamin, D. J., Berger, J.O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Johnson, V.E., et al. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6–10. Doi:10.1038/s41562-017-0189-z

Dearlove, D., & Coomber, S. (1999). *Heart and Soul and Millennial Values.* Skilman, NJ: Blessing/White.

De Bruin, G. P. (2004). Problems with the factor analysis of items: Solutions based on item response theory and item parcelling. *South African Journal of Industrial Psychology*, *30*(4), 16–26. https://doi.org/1.4102/sajip.v30i4.172

Diskiené, D., & Goštautus, V. (2013). A fit between individual and organizational values and its implications for employees' job satisfaction and performance. *Ekonomika, 92*(2), 93-107. Retrieved from http://www.zurnalai.vu.lt/ekonomika/article/viewFile/1412/801.

Erikson, E. (1982). *The life cycle completed*. New York: Norton.

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

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd Ed.). New York: Guilford.

Little, T. D., Rhemtulla, M., Gibson, K., & Schoeman, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*(3), 285–3. https://doi.org/1.1037/a0033266.

Marsh, H. W., Hau, K.T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. (Expanded ed.) Chicago: The University of Chicago Press.

Schwartz, S. (1999). A theory of cultural values and some implications for work. *Applied Psychology: An International Review, 48(*1), 23–47. http://dx.doi.org/10.1111/j.1464-0597.1999.tb00047.x

Strobl, C., Kopf, J., & Zeileis, A. (2011) : A new method for detecting differential item functioning in the Rasch model. *Working Papers in Economics and Statistics*, No. 2011-01. Retrieved from http://hdl.handle.net/10419/73503

Sullivan, W., Sullivan, R., and Bufton, B. (2002) Aligning individual and organisational values to support change. *Journal of Change Management, 2*(3), 247–254. doi: 10.2307/256007

Yves, R. (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/.