Taylor, N. (2008, June 30). Raven's Standard and Advanced Progressive Matrices among Adults in South Africa. *WebPsychEmpiricist*. Retrieved June 30, 2008 from http://wpe.info/papers_table.html.

WPE

WebPsychEmpiricist

Raven's Standard and Advanced Progressive Matrices among Adults in South Africa

Nicola Taylor

Jopie van Rooyen & Partners SA (Pty) Ltd

June 30, 2008

Abstract

Due to the multicultural nature of the South African population and the fact that the country boasts 11 official languages, Raven's Standard (SPM) and Advanced (APM) Progressive Matrices are often used in organisational contexts as measures of cognitive ability. The emergence of the Employment Equity Act (55 of 1998) created a hesitance in the commercial sector with regard to the use of psychological assessments, as the Act clearly stipulates that psychometric assessments may not be used unless they have proven reliability and validity, are not biased against any employee, and can be fairly applied to any employee or group. The non-verbal nature of the SPM and APM lends them to the assumption of fairness, as language ability is excluded from the measurement of cognitive ability. However, questions have arisen as to whether these assessments measure the same construct in different groups, and whether the test is biased against individuals classed as previously disadvantaged through the apartheid system. The present study was undertaken in order to investigate whether Raven's Standard and Advanced Progressive Matrices function similarly for Black and White working adults in the South African context. Item response theory, as conceptualised by the Rasch model (Rasch, 1960) was used to investigate whether or not the tests measure cognitive ability in the two groups in a similar way, and whether or not there is any evidence for bias in either the SPM or APM. Results show there is no such bias.

Address correspondence regarding this article to: Nicola Taylor, P.O. Box 2560, Pinegowrie, 2123 South Africa. Email: nicola@jvrafrica.co.za. An earlier version of this article has been published in Raven & Raven (eds) (2008).

Raven's Standard and Advanced Progressive Matrices among Adults in South Africa

June 30, 2008

Experiment 1: Standard Progressive Matrices (Classic Form)

The first research with the Standard Progressive Matrices (SPM) in South Africa was by Rimoldi in 1945. This provided SPM percentiles for children aged 7 – 14 for each gender (Rimoldi, 1945). Since then, the SPM has proved to work effectively and reliably in the South African context. However, most of the studies have focused on schoolchildren (e.g., Crawford-Nutt, 1976; Owen, 1992), and very few norm tables are available for adult samples in South Africa. Also, only a handful of published studies on the functioning of the Classic SPM items across cultural groups are available in South Africa (e.g., De Bruin, De Bruin, Derckson, & Cilliers-Hartslief, 2005).

Method

Participants

The data used for the analyses to be reported in this chapter were extracted from Jopie van Rooyen & Partners' (JvR) Consulting Services database. They were accumulated during selection exercises carried out for four major clients between 2005 and 2007. The demographics for the Classic SPM study are shown in Table 1. It will be seen that the data were provided by 144 female and 199 male job applicants, of whom 46.9% were Black and 41.8% White. The average age was 33.8 years. For the purposes of the following analysis, only the Black and White applicant groups will be compared, as the Indian and Coloured groups were too small.

Results

Internal consistency.

Although, as has been shown in earlier chapters, it is not entirely appropriate to calculate conventional measures of internal consistency for Item Response Theory (IRT) - based tests, these were generated to meet users' expectations.

Table 2 shows the Cronbach alpha coefficients for each of the five sets of the Classic SPM, as well as that for the test as a whole. The internal consistency of Set A is lower than the others. This is because it consists mainly of easy items, which most adults are likely to get right. Also, as part of the demonstration of the answering procedure required, respondents are given the correct answers to the first two items on Set A, so these were removed from the reliability analysis. Overall, the internal consistency reliability of the SPM is very good.

Descriptive statistics.

Table 3 presents descriptive statistics for the SPM for the entire group. As can be seen by looking at the mean scores for each set, respondents tend to get progressively fewer items correct the further they progress through the test. Overall, respondents got 45 out of the 60 items comprising the Classic SPM correct.

The skewness statistic gives an indication of how easy the respondents found the test. A negative skew means that most of the respondents got fairly high scores on the test. Figure 1 shows that the majority of scores on the SPM were at the higher end of the distribution. The kurtosis statistic is an indication of how sharp the curve of the distribution line is, or how concentrated the scores are. Positive kurtosis indicates a sharper spike in the curve than one would expect from a Gaussian – often misleadingly termed a "normal" – distribution. Figure 1 again illustrates that the scores are fairly highly concentrated around the mean score.

Group comparisons.

Scores on the SPM were compared across gender and ethnic groups using an independent samples t-test. The results of the t-test for the gender groups are presented in Table 4, and the results of the t-test for the ethnic groups are presented in Table 5.

Table 5 shows that, on all Sets, and on total score, the White group on average scored significantly higher than the Black group. This finding may cause some concern at first, but it is important to consider the context in which the test was administered.

Figure 2 shows the separate score distributions for the Black and White applicant groups. From the graph, it can be seen that there was more of a spread of scores for Black applicants than for White applicants in terms of the total SPM score. This may well be due to the pre-selection process used to screen applicants before they reach the assessment phase. Most companies in South Africa are governed by Employment Equity policies, as well as Industry standards, which often predetermine the demographic profile of employees at different levels within the organisation. The shapes of the distribution of SPM scores (as seen in Figure 2) for Black and White applicants show that the distribution of the Black applicant group is more Gaussian than that for the Whites. This may be an indication that, on average, a select group of White applicants of higher ability are invited to participate in the assessment process, whereas the

Black applicants who are invited to participate in the assessment process vary more widely in their level of ability as measured by the Classic SPM.

Differential item functioning.

It has been widely argued that many psychological tests are "biased" against certain, often minority, groups. The meaning and interpretation of this claim has been highly contentious, often involving legal proceedings*. Despite repeated demonstrations (from 1938 onwards) that the RPM items scale in much the same way in most cultural groups (see Raven, 2000, for a review), the charge that it is unfair to certain groups continues to be levelled against it. The basis for this claim is that certain cultural groups "think in different ways" or, are, at the very least, "unfamiliar with the way of thought" required to perform well on the test. Yet, if this were the case, the items would not "scale"; their difficulty indices would be random, or, at the very least, when arranged in order of difficulty, the items would not be in the same order.

Although this charge has repeatedly been shown to be unfounded, a number of tenable arguments and studies are still put forward to support this position. Most of these claims are based on attempts to apply the inappropriate assumptions of Classical Test Theory to a test which, as we have seen in other chapters, was built on, and conforms to, the requirements of, Item Response Theory. Nevertheless, because of the social significance of the claim in Africa, a serious effort was made to address the question (using the latest available techniques) in the course of the present study.

In the IRT literature, the question has been tackled under the rubric of "differential item functioning". Unfortunately, disputes at various conferences indicate that the term is not always understood in the same way and that there are disagreements about the best way of assessing "it".

Our own impression on the issues is that there are two major groups of researchers. One is concerned solely with the question of whether the test as a whole "scales" in the same way among different cultural groups once the item difficulties are established using IRT based procedures … and, in the process, identifying those items which do not. The other group is concerned with whether, in addition, the individual Item Characteristic Curves have the same shape for both groups.

---

*Jensen (1980) offers a fairly thorough discussion of the issues.

As argued by Raven, Prieler and Benesch (2005)*, the latter question can only be investigated if the ICCs are generated using a 3pl model. Since the present investigation was conducted using a Rasch model, often viewed as a 1pl model, this was without the scope of the present study.

The aim of those investigating differential item functioning is to find out whether test takers who have similar knowledge (as determined from total test scores) perform in similar ways on individual test questions regardless of their gender, age, or ethnicity.

The main premise of item response theory is that the higher a person's ability is relative to the difficulty of an item, the higher the probability of a correct response on that item. The item response theory model used to analyse the data in the present sample was the Rasch model, using the WINSTEPS® program version 3.58.1 (Linacre, 2005). The Rasch model is the only measurement model that can transform human data into abstract, equal-interval scales, while maintaining strict objective criteria for the construction of a scale that is separate from the sample distribution (Bond & Fox, 2001). As observed by Linacre (1996), "failure of a data set to fit the Rasch model implies that the data do not support the construction of measures suitable for stable inference" (p. 512).

Using a Rasch analysis means that the estimated item difficulties are theoretically independent of the characteristics of the sample of persons taken from the population of interest. In many statistical approaches, knowledge of the sample distribution is required or assumed. In Rasch, the details of the sample distribution are generally unknown until after the analysis is completed. So, in most cases, the estimated item difficulties are statistically the same when whether the sample as a whole is high or low performing, central or dispersed, unimodal or multimodal, skewed or symmetrical. Of course, this ideal is never achieved to perfection, but it at least allows the researcher to make inferences about the test regardless of the distribution of the sample (J. M. Linacre, personal communication, August 2007).

In order to best illustrate this point, it is useful to examine the person-item map produced by a Rasch analysis. Because item difficulties and person abilities are calculated on the same scale (called "log-odds units" or "logits" in Rasch terminology), it is easy to see the sample

* This reference can be found on WPE's Papers Table. The article was updated in Raven & Raven (eds) (2008).

characteristics compared to the item characteristics. Figure 3 shows the person-item map for the Classic SPM items and the respondents tested for the present study.

In Figure 3, the items appear on the right hand side of the line, distributed according to difficulty. The mean item difficulty for the items on the SPM is 0 logits – negative logits indicate easier items, and positive logits indicate more difficult items. The easiest items are A1, A2, A3, and A4, and the most difficult items are E11 and E12, as one would expect. The applicants appear on the left hand side of the map, distributed according to ability. The mean person ability is 2 logits, which is one standard deviation above the item difficulty. This is an indication that the items are actually too easy for many of the applicants, as can be seen by a number of applicants whose ability lies far above the highest level of ability tapped by the most difficult items, E11 and E12*.

As we have seen, if a test conforms to the Rasch model, the estimated item difficulties should be independent of the ability level of the sample which provided the basic data used to calculate them. Because the significance of this statement will be lost on many readers, attention may be drawn to just how starkly it differs from the situation that prevails when traditional indices of difficulty are employed. These show the proportion of the population tested who choose the correct answer to a given question. Clearly, the proportion of a low ability sample which gets an item right is very much lower than the proportion of a high ability sample which gets it right. While these item difficulties may rank-order the items in the same way, the difficulty indices themselves are anything but identical.

When using the Rasch model to seek evidence of differential item functioning between two groups, the question is, therefore, “To what extent is it true that the actual item difficulties, expressed in logits, are different for the samples which yielded the data on which they are based?”

The Rasch item difficulties assessed separately for the Black and White groups have therefore been plotted side by side in Figure 4 with a view to identifying those items which may not be functioning in the same way in both groups. In principle, these can be identified using a t-test … although, by definition, since there are 60 items in the SPM a proportion will have significantly different difficulty indices for purely statistical reasons.

---

* This, of course, is the reason why first the APM, and, later, the SPM ***Plus*** tests were developed.

It is clear that, in general, the items are functioning in very similar ways: Only four items could be flagged as having statistically significant difficulty indices (i.e. items B7, B8, C4 [which the White respondents found relatively more difficult], and E5 [which the Black respondents found relatively more difficult]). More detailed analyses would be required to determine whether these statistically significant differences reflect meaningful differences in psychological functioning.

The overall correlation between the Rasch-based item difficulty indices established separately in the Black and White groups was 0.97, again indicating that, despite the overall difference in average scores between the two groups, the test is functioning in an almost identical way within the two groups. It follows that common-sense-based "explanations" of the difference in mean score between the groups along the lines that "the test is unfamiliar to the way of thought of the Black group" do not hold up. The test works, and works in the same way, in both groups. Explanations of the difference must be sought elsewhere.

The overall functioning of the test.

By now, we have established that the SPM is functioning in a similar way at the item level in both groups. It remains to ask whether the test as a whole functions in the same way for both groups.

This question was also encountered in the investigation reported in the paper *Lessons Learnt Whilst Developing a Romanian version of the MHV* (Raven, 2006)[*]. There, it was shown that one could in no sense assume that the overall Test Characteristic and Test Information Function curves developed in different settings would be similar, still less that these curves would, or should, conform to a Gaussian ogive[†].

In the light of these earlier discussions, the similarity between the Test Characteristic Curves calculated separately for the Black and White groups shown in Figure 5 is striking indeed.

Perhaps the most powerful statement that can be made on the basis of these results is that it would appear that Blacks and Whites who have the same ability as determined by their scores

[*] This article is also included in Raven & Raven (eds) (2008).

[†] Some readers may find it useful to be reminded that the Test Characteristic Curve is generated by summing the information contained in the individual Item Characteristic Curves.

on the latent continuum have the same raw scores. The test is *not* "biased against them" because of their ethnicity.

Discussion

The results of the analysis of the SPM data show that Black and White applicants of the same ability are equally likely to achieve the same score on the Raven's Standard Progressive Matrices in a selection context. The score differences obtained using Classical Test Theory methods are not a result of the differential functioning of the test, but more likely to be due to the composition of the sample. The SPM is a reliable measure of eductive ability, and the results of the present study are consistent with previous research.

Experiment 2: The Advanced Progressive Matrices

. Raven's Advanced Progressive Matrices (APM) was developed to differentiate between people of superior intellectual ability (Raven, Raven & Court, 1998). It is generally used in selecting employees for high-level technical or managerial positions. Research done on the APM in South Africa is limited, although some studies have been conducted in the South African National Defence Force (Muller & Schepers, 2003).

Method

Materials. The APM has two components: Set I consists of 12 items that are often used as practice, as they provide training in the method of thinking required to complete Set II effectively. Set II consists of 36 items of increasing difficulty and constitutes the main part of the test (Raven, et al., 1998).

Participants.

The demographics for the APM group are shown in Table 6. 32 women and 158 men were involved, of whom 35.1% were Black and 60.2% White. The average age was 37.1 years. For the purposes of the following analysis, only the Black and White groups will be compared.

Results

Internal consistency reliability.

The internal consistency reliability for Set I of the APM (12 items) is 0.57, which is low. The reason for this low reliability is probably due to the low difficulty level of the items in the first set, which is usually used as a training set, where candidates can clarify any of the items of which they are unsure. The internal consistency reliability for the Set II of the APM (36 items) is 0.89, which can be described as acceptable.

Descriptive statistics.

Table 7 presents the descriptive statistics on the APM for the entire group. As can be seen by looking at the mean scores for each set, respondents tend to get most of the Set I items correct. Overall, respondents got an average of 23 out of 36 items correct on Set II. From this point forward, only Set II scores will be included in analyses.

The skewness and kurtosis statistics for the Set II are very close to 0, indicating that the scores are distributed in a Gaussian curve. Figure 6 shows that the distribution of scores in graphic format.

Group comparisons.

Scores on the APM were compared across gender groups and ethnic groups using an independent samples t-test. The results showed no significant difference between men (Mean = 22.56, SD = 7.23) and women (Mean = 23.41, SD = 5.16) on the Set II APM [$t(188) = -0.632$, $p = 0.528$]. However, there was a significant difference between Black applicants (Mean = 19.01, SD = 7.01) and White applicants (Mean = 24.83, SD = 6.08) [$t(180) = -5.872$, $p = 0.000$].

The separate score distributions in Figure 7 show a similar pattern to the Classic SPM distributions across ethnic groups. Again, it is highly likely here that the difference in score distributions is due to the same selection practices employed in the use of the SPM.

Differential item functioning.

In Figure 8, the items appear on the right hand side of the line, distributed according to difficulty. The mean item difficulty for the items on the APM Set II is 0 logits. The easiest items are 1, 2, and 3 and the most difficult items are 32 and 36, as one would expect. The applicants appear on the left hand side of the map, distributed according to ability. The mean person ability is just less than 1 logit, which is within one standard deviation above the mean item difficulty. This is an indication that, for the most part, the items are fairly well matched to the ability of applicants. All the applicants are more able than level tapped by the first three items, although there are some applicants whose ability lies above the level tapped by the most difficult item, number 36.

Figure 9 plots the Rasch item difficulties for Black and White applicants on Set II of the APM. The graph shows slightly different patterns of item difficulty for the two groups, but they show the same tendency that the items get more difficult as they progress through the test. Only five items could be flagged as possibly reflecting Differential Item Functioning. These were items 9 and 10, (which the Black applicants found more difficult), and 17, 19 and 28 (which the White applicants found more difficult). Although it appears that there may be differences in item difficulty for other items in Set II, they are not significant. The correlation between the item difficulties (in logits) determined separately among Black and White respondents was 0.93, which indicates that, despite the overall difference in average scores between the two groups, the test is functioning in an almost identical way within the two groups.

The Test Characteristic Curve (TCC) (calculated by cumulating the individual Item Characteristic Curves) gives an indication of what the expected raw score that someone having a given level of ability would be expected to attain. The TCC's for the Black and White applicants are shown in Figure 10. The resulting TCC's show that the curves are very similar, indicating that individuals of the same ability from either group are equally likely to obtain similar raw scores on the APM Set II.

Discussion

The results of the analysis of the APM data show that Black and White applicants of the same ability are likely to achieve a similar score on the Raven's Advanced Progressive Matrices in a selection context. The raw score differences are again not a result of the differential functioning of the test, rather the composition of the sample.

Some Conclusions

The results of the analysis of both the Classic SPM and APM data indicate that the claim that Black and White South Africans perform differently on these two tests is unsubstantiated. Regardless of cultural group, individuals of a certain ability level should be able to obtain the same raw score as others of the same ability.

The finding that the Black group did score lower on average than the White group is most likely a function of the sample characteristics. A larger, more representative sample is required before inferences are made as to why the nature of the samples differs.

Tables

Table 1

Classic Standard Progressive Matrices

Demographics of the population studied

	Group	**N**	**%**
Gender	Women	144	33.8
	Men	199	46.7
	Unspecified	83	19.5
Total		426	100.0
Ethnicity	Black	200	46.9
	White	178	41.8
	Indian	32	7.5
	Coloured	4	0.9
	Unspecified	12	2.8

Table 2

Classic Standard Progressive Matrices

Internal consistency

Set	**Number of items**	**Cronbach's alpha**
Set A	12	0.75
Set B	12	0.87
Set C	12	0.85
Set D	12	0.86
Set E	12	0.86
Total Test	60	0.96

Table 3

Classic Standard Progressive Matrices

Overall Descriptive Statistics by Set

Set	N	Min	Max	Mean	SD	Skewness	Kurtosis
Set A	426	4	12	10.95	1.59	-2.08	4.31
Set B	426	0	12	10.01	2.74	-1.84	2.60
Set C	426	0	12	8.80	2.93	-1.26	1.14
Set D	426	0	12	9.17	2.89	-1.75	2.64
Set E	426	0	12	5.73	3.40	.06	-1.12
Total score	426	6	60	44.65	11.94	-1.35	1.40

Table 4
Classic Standard Progressive Matrices
Mean Scores by Gender

	Men (N = 199)		**Women (N = 144)**			
Set	**Mean**	**SD**	**Mean**	**SD**	**t**	**P**
Set A	11.01	1.58	10.97	1.48	0.195	.846
Set B	9.90	2.92	10.07	2.55	-0.544	.587
Set C	8.87	3.05	8.60	2.92	0.845	.399
Set D	8.95	3.08	9.35	2.80	-1.209	.227
Set E	5.95	3.67	5.47	3.05	1.307	.192
Total score	44.69	12.64	44.45	11.28	0.183	.855

Note. The results of the t-test across gender groups show that there are no significant differences on any of the Sets of the Classic SPM, or on the total score.

Table 5
Classic Standard Progressive Matrices
Mean scores by ethnicity

	Black (N = 200)		**White (N = 178)**			
Set	**Mean**	**SD**	**Mean**	**SD**	**t**	**P**
Set A	10.61	1.86	11.37	.98	-4.881	.000
Set B	9.45	3.17	10.62	1.98	-4.232	.000
Set C	8.12	3.09	9.50	2.60	-4.664	.000
Set D	8.43	3.28	9.89	2.19	-5.041	.000
Set E	4.60	3.16	6.84	3.33	-6.714	.000
Total score	41.20	13.06	48.21	9.33	-5.940	.000

Table 6
Advanced Progressive Matrices.
Demographics of Population Studied.

	Group	N	%
Gender	Women	32	16.8
	Men	158	82.7
	Unspecified	1	0.5
Total		191	100.0
Ethnicity	Black	67	35.1
	Indian	8	4.2
	White	115	60.2
	Unspecified	1	0.5
Education	Grade 10	1	0.5
	Grade 12	11	5.8
	Higher diploma, National diploma, National certificate	80	41.9
	First degree, Honours degree	79	41.4
	Masters degree, Professional qualification	11	5.8
	Unspecified	9	4.7

Table 7
Advanced Progressive Matrices.
Descriptive statistics by Set

Set	N	Min	Max	Mean	SD	Skewness	Kurtosis
Set I	191	3	12	10.48	1.57	-1.73	4.81
Set II	191	1	36	22.70	6.90	-.36	-.31

Figures

Figure 1
Classic Standard Progressive Matrices
Histogram of Total Scores

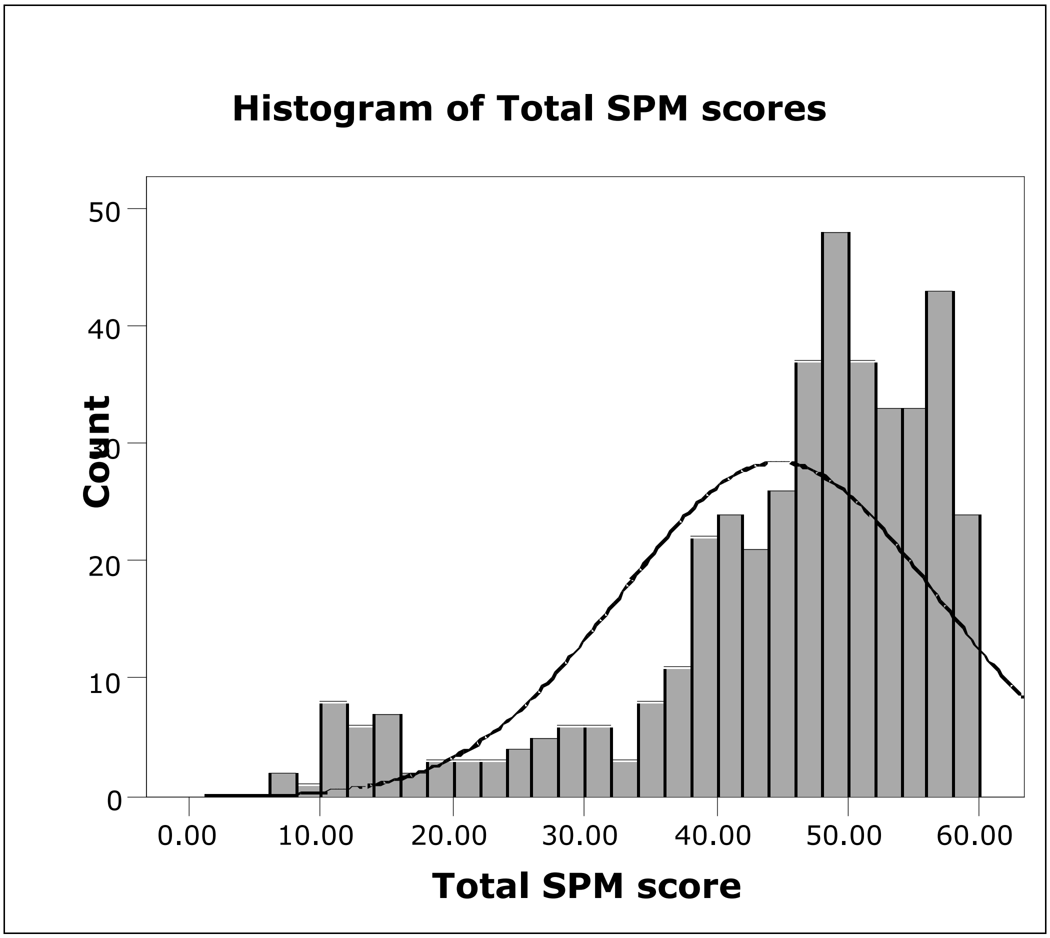


Figure 2
Classic Standard Progressive Matrices
Distribution of scores for Black and White applicants

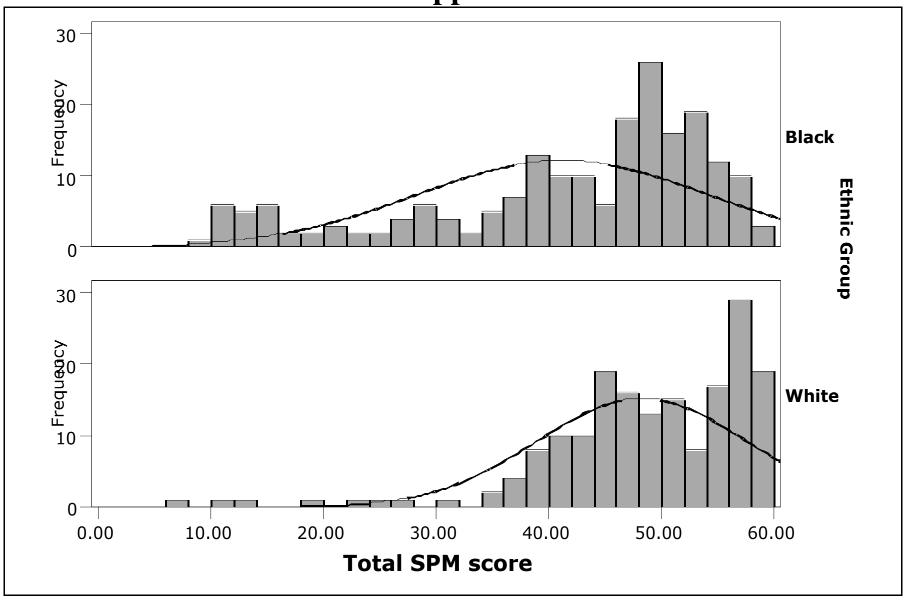


Figure 3
Classic Standard Progressive Matrices
Person-item map

```
        PERSONS MAP OF ITEMS
               <more>|<rare>
    7             .##  +
                       |
                       |
                 .###  |
                       |
    6                 T+
                       |
                       |
               ######  |
                       |  E11     E12
    5                  +
          .##########  |
                       |
          ###########  |T
                       |  E10
    4     .########## S+
               ######  |  C12
                       |  D12
              .######  |
           ##########  |
    3     ###########  +
             .#######  |  D11     E8      E9
        .############  |
         .###########  |
        .############ M|S E6      E7
    2          ######  +  C11     E4
        #############  |
                .####  |  B12     C10     C8      E5
               ######  |
         ############  |  A12
    1          .#####  +
            .########  |  C6      D10     E2
                 .###  |  B7      B8      E3
                  ###  |  A11     B11
                   .# S|  B9      C9      D7      D9      E1
    0             ###  +M C4      D8
                  ###  |
                   .#  |  B6      C5      C7      D6
                   ##  |  B10     D4
                   ##  |
   -1               .  +  A7      B5
                   .#  |  A10     B4      C3      D3      D5
                   .#  |  D2
                    .  |  C2
                    . T|  B3      C1
   -2            .###  +  A8      A9      D1
                    #  |S
                   ##  |  B2
                  ###  |
                    #  |
   -3               .  +
                       |  A5      A6      B1
                       |  A1
                    #  |
                       |
   -4                  +
                       |  A2
                       |T A3
                       |  A4
                       |
   -5                  +
               <less>|<frequ>
 EACH '#' IS 2.
```

Figure 4

Classic Standard Progressive Matrices

Plot of item difficulties scaled separately for Black and White respondents.

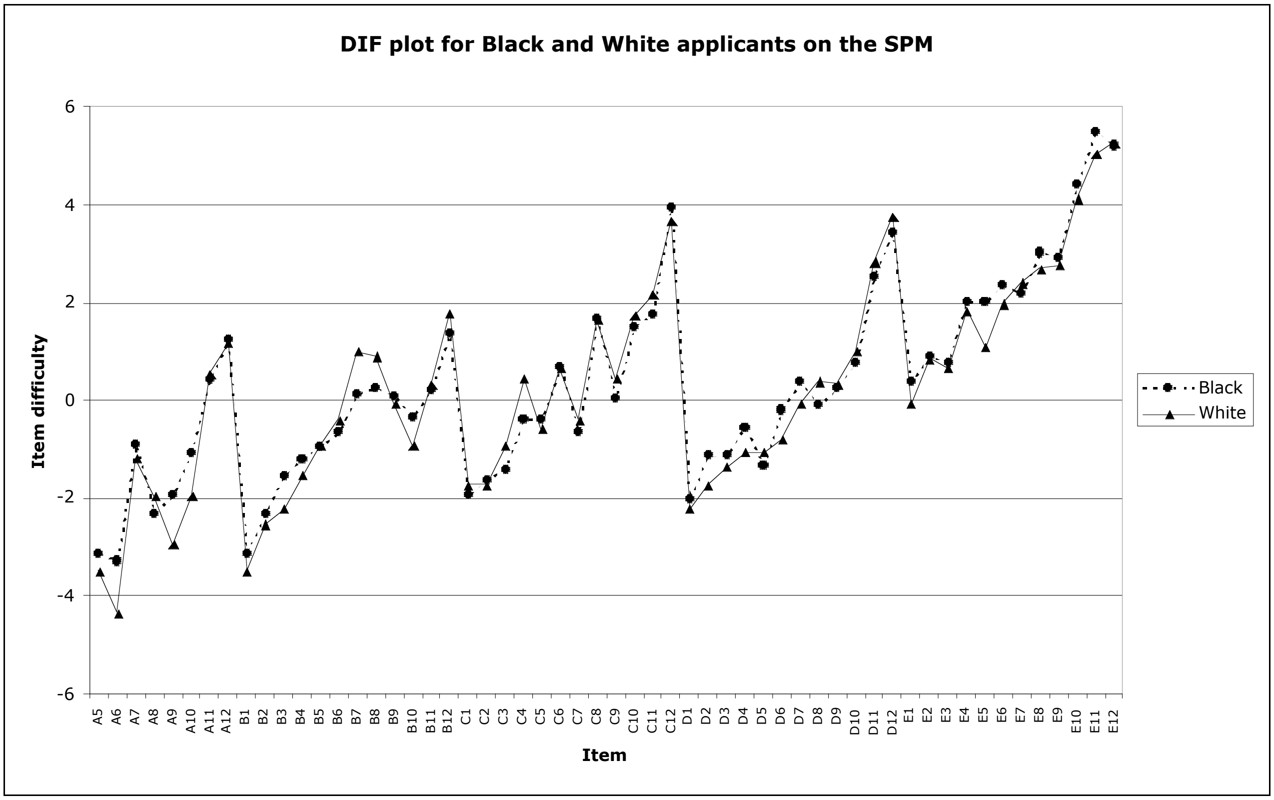


Figure 5

Classic Standard Progressive Matrices

Test characteristic curves for Black and White Respondents

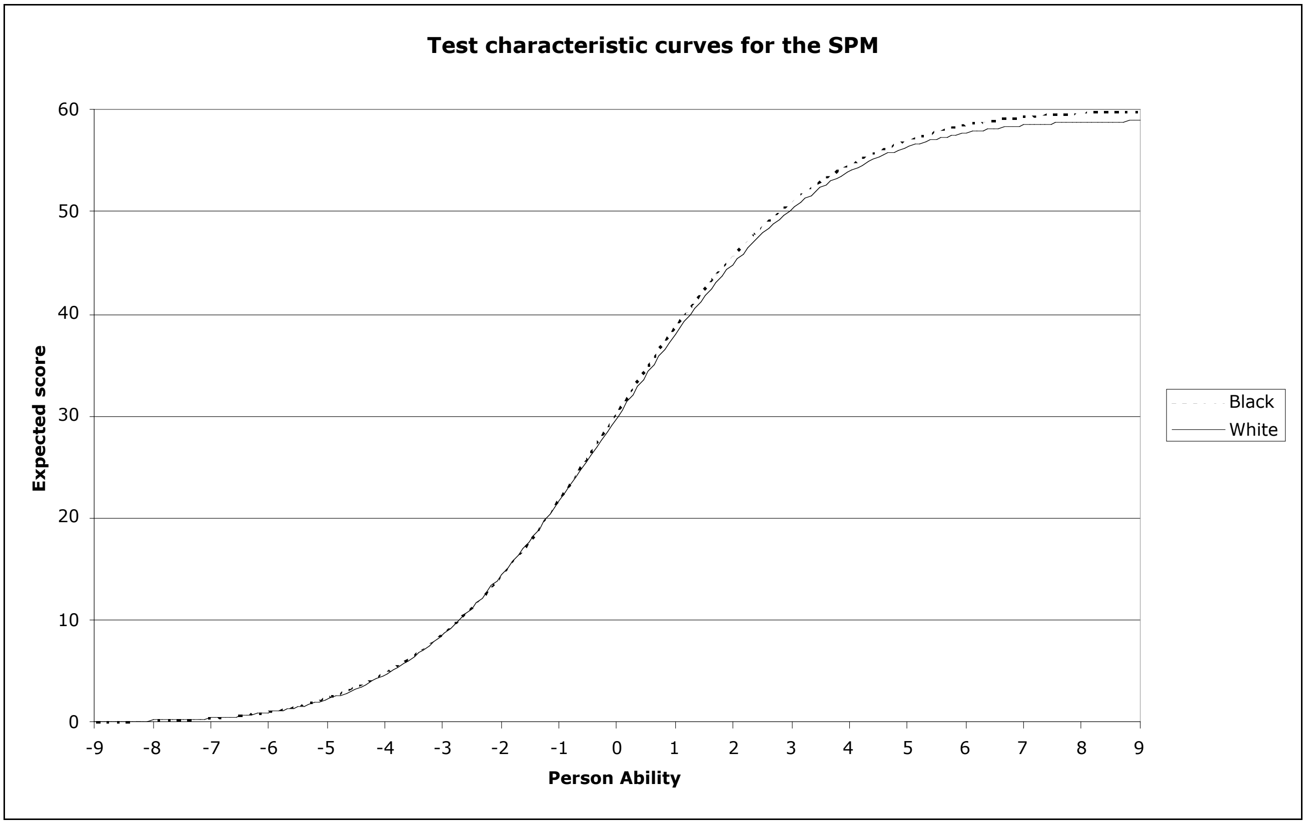


Figure 6
Advanced Progressive Matrices, Set II.
Overall Score Distribution

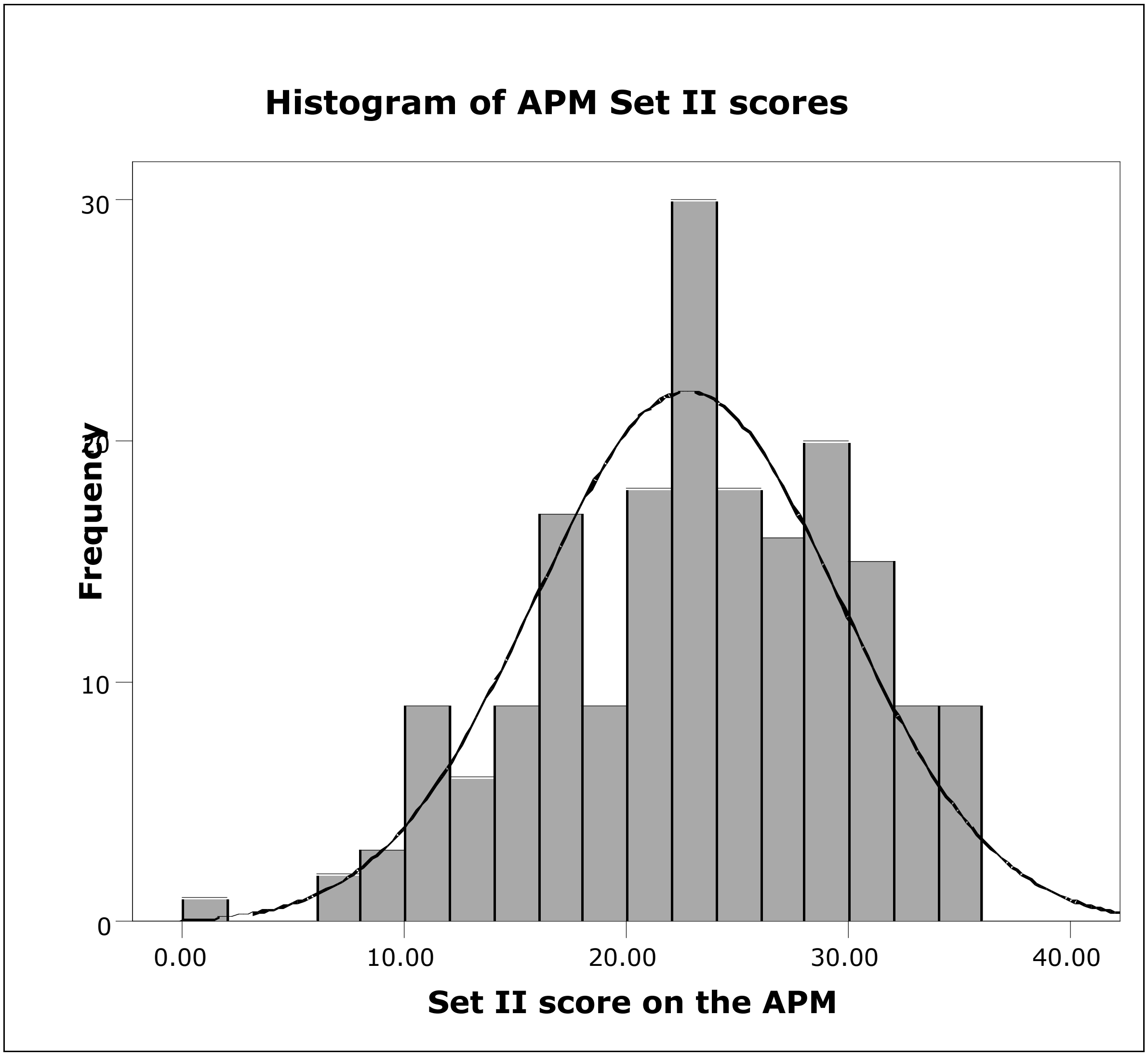


Figure 7
Advanced Progressive Matrices, Set II.
South African Adult Study.
Score Distributions by Ethnicity

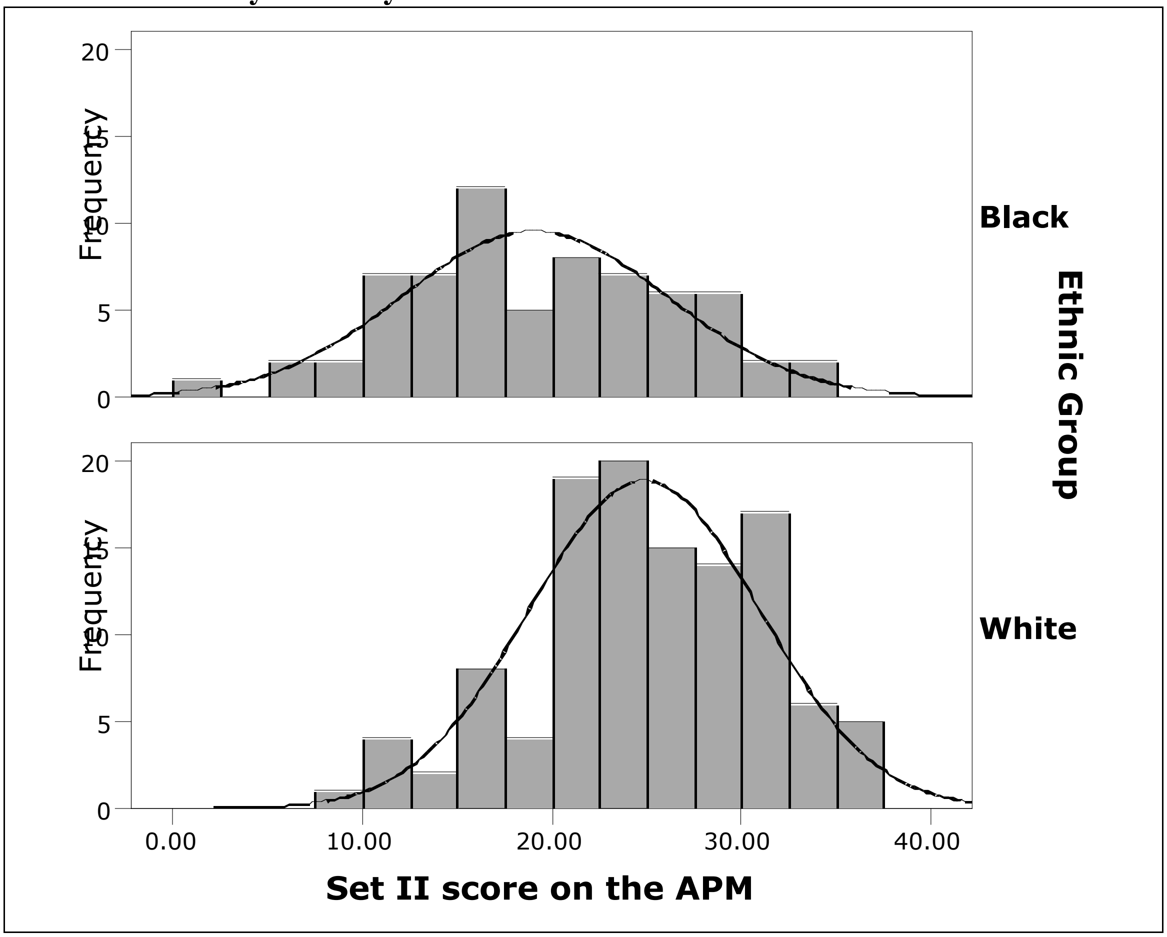


Figure 8

Advanced Progressive Matrices, Set II.

Person-item map

```
         PERSONS MAP OF ITEMS
                <more>|<rare>
    5                .  +
                        |
                        |
               ##       |
                        |
                        |
    4                   +
                        |
               ##  T|
                        |   s36
                        |
               ##       |
    3                   +T
              .##       |
                        |
             ####       |
                        |
             .### S|    s32
    2       #####       +
            #####       |
                        |   s27   s29   s31   s35
            .####       |S  s28   s30   s34
             .###       |   s33
             .###       |   s24
    1      .#####       +   s25   s26
        ######### M|
           ######       |   s22   s23
           .#####       |
             .###       |   s18
             .###       |   s21
    0               #   +M  s19   s20
            .####       |
             ####       |
               .# S|    s10   s16   s17
              ###       |   s13
               ##       |
   -1               #   +   s14
               ##       |   s11   s12   s15   s4
              .##       |   s5
               .#       |S  s8    s9
                        |   s6
                   T|   s7
   -2               #   +
                        |   s3
                        |   s2
                        |
                        |   s1
                        |
   -3                   +T
                        |
                        |
                        |
                        |
                        |
   -4                   +
                        |
                .       |
                        |
                        |
                        |
   -5                   +
                <less>|<frequ>
 EACH '#' IS 2.
```

Figure 9
Advanced Progressive Matrices, Set II
Plot of item difficulties calculated separately for Black and White respondents.

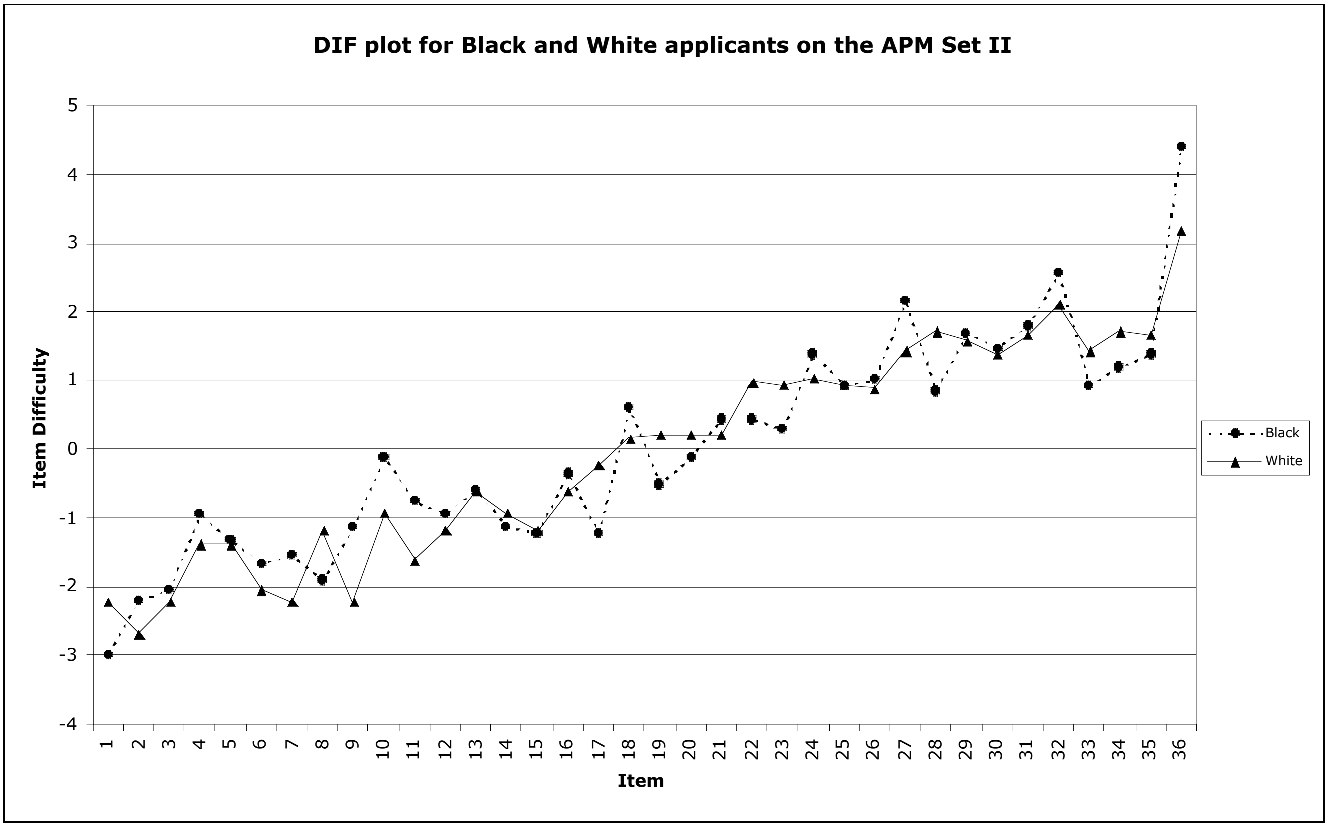
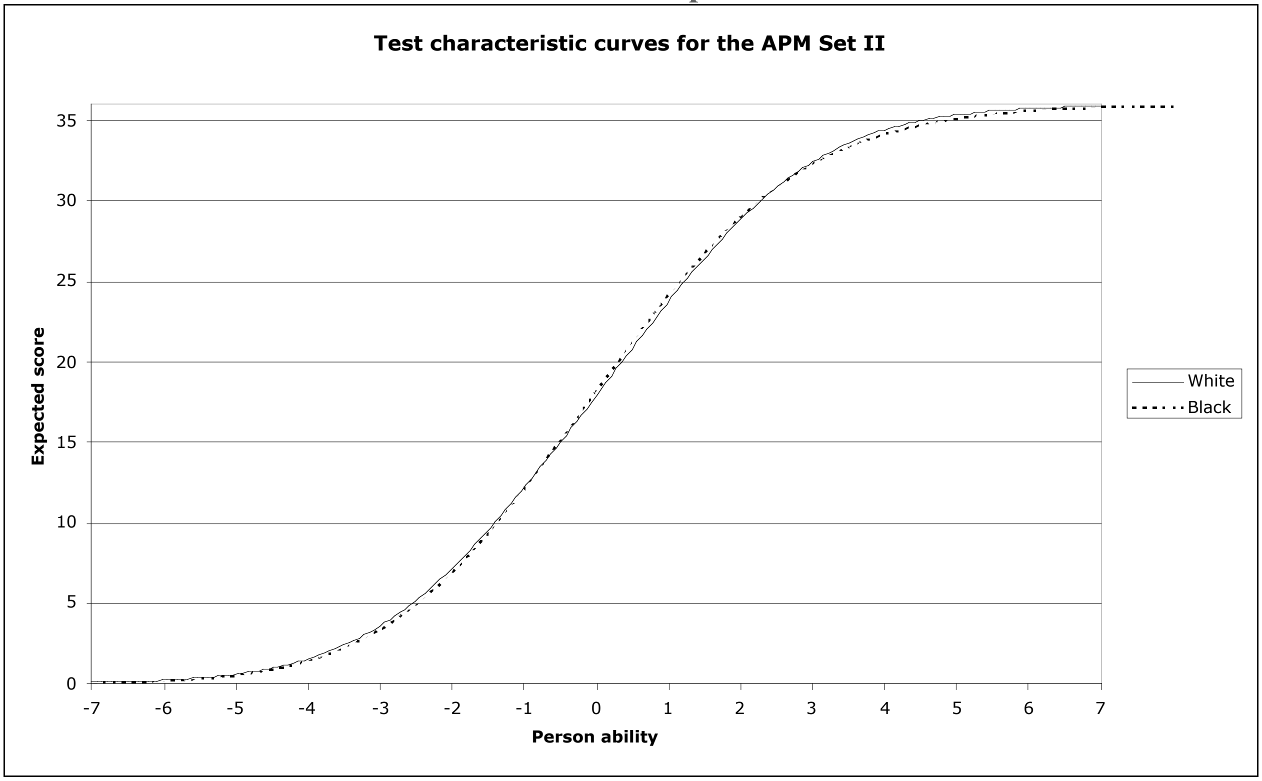


Figure 10
Advanced Progressive Matrices, Set II
Test characteristic curves for Black and White respondents.



References

Bond, T.G. & Fox, C.M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Crawford-Nutt, D.H. (1976). Are Black scores on Raven's Progressive Matrices an artifact of method of test presentation? *Psychologia Africana, 16,* 201-206.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-344.

De Bruin, K., De Bruin, G.P., Derckson, S., & Cilliers-Hartslief, M. (2005). Predictive validity of general intelligence and Big Five measures for adult basic education and training outcomes. *South African Journal of Psychology, 35,* 46-57.

Jensen, A. R. (1980). *Bias in Mental Testing*. New York: Free Press.

Linacre J.M. (1996). The Rasch Model cannot be "Disproved"! *Rasch Measurement Transactions,* 10, 512-514.

Linacre J.M. (2005). *WINSTEPS® Rasch Measurement*. Available from www.winstepscom.

Muller, J., & Schepers, J. (2003). The predictive validity of the selection battery used for junior leader training within the South African national defence force. *SA Journal of Industrial Psychology, 29,* 87-98.

Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences, 13* (2), 149-159.

Rasch, G. (1960). *Probalistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Instituut.

Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology, 41*, 1-48.

Raven, J., Prieler, J., & Benesch, M. (2005, June 15). A Replication and Extension of the Item-Analysis of the Standard Progressive Matrices Plus, Together With a Comparison of the Results of Applying Three Variants of Item Response Theory. *WebPsychEmpiricist*. Retrieved June 15, 2005 from http://wpe.info/papers_table.html.

Raven, J., & Raven, C. J. (Eds.) (2008). *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*. Unionville, New York: Royal

Fireworks Press; Edinburgh, Scotland: Competency Motivation Project; Budapest, Hungary: EDGE 2000; Cluj Napoca, Romania: Romanian Psychological Testing Services SRL.

Raven, J., Raven, J. C., & Court, J. H. (1998, updated 2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. *Section 1: General Overview*. San Antonio, TX: Harcourt Assessment.

Raven, J., Raven, J. C., & Court, J. H. (2000, updated 2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. *Section 3: The Standard Progressive Matrices, Including the Parallel and Plus Versions*. San Antonio, TX: Harcourt Assessment.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. *Section 4: The Advanced Progressive Matrices*. San Antonio, TX: Harcourt Assessment.

Rimoldi, H.J.A. (1948). A note on Raven's progressive matrices test. *Educational and Psychological Measurement, 8,* 347-352.