Available online at www.sciencedirect.com
SCIENCE DIRECT®
Intelligence 32 (2004) 411–424
ELSEVIER
INTELLIGENCE

Sex differences on three factors identified in Raven's Standard Progressive Matrices

Richard Lynn[a,*], Jüri Allik[b], Paul Irwing[c]

[a]University of Ulster, Coleraine, Northern Ireland, United Kingdom
[b]University of Tartu, Estonia
[c]University of Manchester Institute of Science and Technology, United Kingdom

Received 24 November 2003; received in revised form 31 May 2004; accepted 8 June 2004
Available online 30 July 2004

Abstract

Raven's Standard Progressive Matrices (SPM) was administered to a sample of 2735 12- to 18-year-olds in Estonia. Both a scree test and the consistent Akaike information criterion (CAIC) indicated the presence of three significant factors. Exploratory and confirmatory factor analysis showed the loadings of the items on the three factors, which were identified as the gestalt continuation found by Van der Ven and Ellis [*Pers. Individ. Differ.* 29 (2000) 45], verbal–analytic reasoning and visuospatial ability. Further analysis of the three factors showed a higher order factor identifiable as *g*. Examination of age by sex differences showed that on all four factors girls performed better than boys at the age of 12, there was no sex difference at age 14, while boys performed better than girls at the age of 17, although not significantly on visuospatial ability.

Keywords: Intelligence; Progressive matrices; Sex differences

1. Introduction

Raven's Standard Progressive Matrices (SPM) is widely used as a test of nonverbal reasoning ability. This ability is generally identified with fluid intelligence and with *g* (e.g., Carroll, 1993; Jensen, 1998;

* Corresponding author. Tel.: +44 1275 392 092.
E-mail address: lynnr540@aol.com (R. Lynn).

0160-2896/$ - see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.intell.2004.06.007

McGrew & Flanagan, 1998). For more than half a century, there has been a debate over whether the SPM is a pure test of *g* or whether it also measures some kind of visualization or spatial ability. The test was constructed by Raven (1939) to be a pure measure of *g* and was accepted as such by Spearman (Spearman & Wynn-Jones, 1951). This position was endorsed by Emmett (1949) based on a factor analysis of the items of the SPM in a sample of 11-year-olds. In recent times, Jensen (1998, p. 541) has contended that "the total variance of Raven scores in fact comprises virtually nothing besides *g* and random measurement error," and Raven, Raven, and Court (2000, p. 34) state that "The Progressive Matrices has been described as one of the purest and best measures of *g* or general intellectual functioning." The issue of whether total scores on the Progressive Matrices can be equated to *g* is important for a second reason. Two of the present authors have recently completed a meta-analysis investigating sex differences on the Progressive Matrices and found that among adults men obtain higher average means than women by about 5 IQ points (Lynn and Irwing, in press). Hence, if the Progressive Matrices is a pure measure of *g*, it would appear that men have higher *g* than women, contrary to received wisdom expressed by numerous scholars including Jensen (1998) and Mackintosh (1996).

However, it has not been universally accepted that the Progressive Matrices is a pure measure of reasoning ability and *g*. A number of scholars have contended that the Progressive Matrices is largely a measure of *g* but also contains a small or fairly small visualization or spatial factor. These include Adcock (1948), Banks (1949), Keir (1949), Gabriel (1954), and Vernon (1950) who have contended that the SPM is largely a pure measure of *g* but also contains a small spatial ability factor. Gustaffson (1984, 1988) concluded that the SPM contains a reasoning factor and a further factor that he designated cognition of figural relations. Hertzog and Carter (1988) have contended that the SPM contains two factors they designate verbal intelligence and spatial visualization. Van der Ven and Ellis (2000) concluded that the SPM contains two significant factors which they identified as (1) gestalt continuation present in early items for which "the correct solution must be found according to some Gestalt continuation rule"; this factor appears to be the same or closely similar to Gustaffson's cognition of figural relations; and (2) analogical reasoning, present in most of the later items, for the solution of which "the subject should deduce, by means of analogical reasoning, that a certain change in the transition from the first element in a row to the next element in a row must be repeated in the following row." They also found three further factors that they identified as "lack of resistance to perceptual distractors" present in five items in the C section of the test; a "coping" present in five items in the E section of the test; and an unidentifiable fifth factor also present in the same five E items.

The same problems of the number of factors and their identification have been addressed in several studies of the Advanced Progressive Matrices (APM), a more difficult version of the SPM. These have also reached conflicting conclusions. Alderton and Larson (1990) and Arthur and Woehr (1993) concluded that the APM is solely a measure of *g*. But Dillon, Pohlmann, and Lohman (1981) concluded that the APM contains two factors that they identified as "pattern addition/subtraction" factor and "detection of pattern progression." Lim (1994) in a study of 15-year-old boys concluded that the APM is a pure measure of reasoning ability for boys but that for girls it also measures spatial ability. DeShon, Chan, and Weissbein (1995) found two factors that they identified as verbal–analytic and visuospatial abilities. Colom and Garcia-Lopez (2002) consider that the APM measures reasoning and spatial abilities.

In this paper, we address the disputed issue of whether the Progressive Matrices is a pure measure of reasoning ability and *g* or of whether it contains items measuring some kind of visualization or spatial factor in addition to reasoning. This issue is not easily resolved by the usual method for ascertaining the numbers of factors in a data set by a principal components analysis followed by a scree test to determine the number of factors, the identification of the first principal component as the general factor, and a varimax rotation of the number of significant factors. The validity of this procedure, as generally implemented, has been questioned. It has been contended that the use of dichotomous items generally attenuates the magnitude of Pearson's coefficient of correlation and may lead to artefacts in which items of similar difficulty form spurious factors (Gorsuch, 1983; Kim & Mueller, 1978). Van der Ven and Ellis (2000) have attempted to overcome these objections by the use of Rasch analysis, but we do not think this is the best method of dealing with the problem because this type of analysis assumes unidimensionality and the problem is to determine whether unidimensionality is present (Van der Linden & Hambledon, 1997). Rasch models make two further assumptions, which arguably render them inappropriate in the present instance. Neither the assumption that all items have equal discriminating power nor that guessing is minimal seem compatible with the published literature on the SPM (Hambledon & Swaminathan, 1985). For these reasons, we have preferred to use two of the several methods of exploratory and confirmatory factor analysis that have been developed for the analysis of dichotomous items and are presented by Christoffersson (1975), Muthén (1984), Gorsuch (1997), and Panter, Swygert, Dahlstrom, and Tanaka (1997).

2. Method

The SPM Test was administered in Estonia to a representative sample of 2735 adolescents whose ages ranged from 12 to 18 years (with the exception of 46 subjects who fell out from this range). The sample was drawn from 27 Estonian-speaking public secondary schools and gymnasiums from different regions of Estonia, covering all 15 Estonian counties, the capital and largest city, Tallinn, several smaller cities (Tartu, Pärnu, Kohta-Järve, etc.), small towns, and rural areas. Since boys and girls attend mixed secondary schools in Estonia, there is no difference in the socioeconomic status between their families. Data were collected in 2001. The test was administered without any time limits. The 60 items in the test were initially analyzed using the Mplus (Muthén & Muthén, 1998) program for exploratory factor analysis of categorical data. Then, in order to provide a more rigorous test of the factor structure, we used confirmatory factor analyses as implemented in Lisrel 8.30 (Bollen & Lennox, 1991).

3. Results

3.1. Number of factors

Fig. 1 shows a scree plot of the eigenvalues of a principal component analysis. There were 11 factors with eigenvalues greater than unity. The parallel analysis indicated the presence of seven factors, which appears to be over inclusive. Inspection of the scree plot showed a clear discontinuity between Factors 3 and 4, with a continuous pattern of scree after the third factor. Simulation has shown that the scree plot is a consistently good indicator of the number of factors (Zwick & Velicer, 1986) and hence a three-factor solution is indicated.

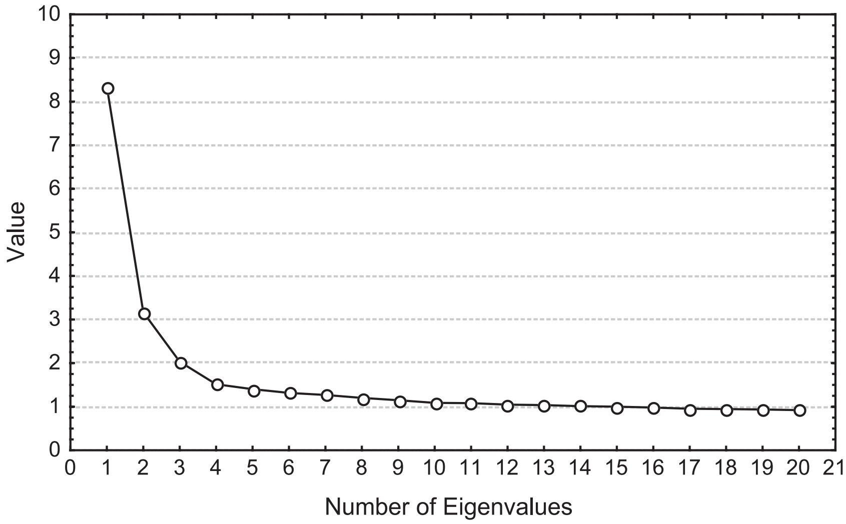

Fig. 1. The scree plot of eigenvalues of the principal component analysis.

3.2. Exploratory factor analyses

The case per item ratio in this study was over 45, which is usually more than sufficient for the stability of factors (Guadagnoli & Velicer, 1988). Traditional methods of factor analysis applied to item-level covariances are questionable for the reasons given by Gorsuch (1983), but several methods of factor analysis have been developed for item level analysis (Christoffersson, 1975; Muthén, 1984; Gorsuch, 1997; Panter et al., 1997). We initially analyzed the data for one-, two-, three-, and four-factor solutions using a mixture of maximum likelihood and weighted least squares exploratory factor analysis for categorical data as implemented in Mplus. These solutions provided the basis for subsequent confirmatory factor analyses.

3.3. Confirmatory factor analyses

In order to verify the factor solution obtained in the exploratory factor analysis, all models were independently tested using LISREL 8.30, with polychoric correlations and maximum likelihood estimation. This was to test the method invariance of the solutions and to provide a more rigorous test for the number of factors (Bollen & Lennox, 1991). Since much of the data provided only ordinal levels of measurement, a weighted least squares analysis was indicated (Jöreskog, 1993). However, responses to some items were too skewed to provide an accurate estimate for the asymptotic covariance matrix. In such circumstances, simulation has shown that maximum likelihood provides the best parameter estimates (Hoogland & Boomsma, 1998).

In order to further resolve the issue of how many latent variables underlie responses to the SPM, we used confirmatory factor analysis to test, one-, two-, three-, and four-factor solutions.[1] Because of the

[1] Although weighted least squares exploratory factor analysis is the technically correct method for categorical data, the four-factor exploratory solution using weighted least squares did not converge when subject to confirmatory analysis. Therefore, in order to provide a standard basis for comparison, the confirmatory factor models used to determine the number of factors were all based on maximum likelihood exploratory factor analyses, all of which provided satisfactory confirmatory models. In the confirmatory analyses, all exploratory loadings <.10 were set at zero.

Table 1
Fit statistics for one-, two-, three-, and four-factor solutions to the SPM based on maximum likelihood and weighted least squares exploratory analyses

Model	SRMR	RMSEA	CAIC
Maximum likelihood-based solutions			
One factor	.049	.059	14,174.0
Two factor	.046	.049	10,738.7
Three factor	.041	.046	9503.8
Four factor	.064	.048	10,332.4
Solution based on weighted least squares exploratory factor analysis			
Three factor	.042	.046	9665.4

highly skewed response distribution of some items, the 60-item correlation matrix was not positive definite. We obtained satisfactory solutions by applying the ridge option to a 52-item correlation matrix. As the four different factor models were nonnested, that model with the minimum consistent Akaike information criterion (CAIC) was selected, following the recommendations of Jöreskog (1993) and Bozdogan (1987). The CAIC for the one-, two-, three-, and four-factor models were 14174.9, 10738.7, 9503.8, and 10332.4, respectively, which points to the optimality of the three-factor solution. Thus, both the scree plot and CAIC support a three-factor solution.

In order to assess the absolute fit of models, we used the standardized root mean square residual (SRMR) and the root mean square error of approximation (RMSEA). We adopted cutoff points of ≤0.05 for both the SRMR (Spence, 1997) and the RMSEA (Browne & Cudeck, 1993). This approach to evaluating fit conforms to recent recommendations based on Monte Carlo simulation (Hu & Bentler, 1999; Marsh, Balla, & Hau, 1996). The relevant fit statistics are reported in Table 1. The one-factor solution does not provide a fit since the RMSEA is in excess of the selected cutoff criterion. Similarly, the four-factor solution does not meet the fit criterion for the SRMR. Both the two- and three-factor solutions meet the specified fit criteria in absolute terms. However, on both the absolute criteria and the CAIC, the three-factor solution clearly represents superior fit. The loadings of the items on the three factors in the three-factor solutions obtained in the confirmatory analyses are shown in Table 2.

3.4. Hierarchical factor analysis

Although three primary factors are indicated by our analyses, given the widespread view that the SPM is an almost pure measure of *g* (Jensen, 1998; Raven et al., 2000), it would be predicted that the three primary factors should all load highly on a second order factor of *g*. Consequently, we tested a model in which all three primary factors were allowed to load on a second order factor. The results are shown in Table 2. This gives the loadings of the items on the three factors in the three-factor solutions obtained in the exploratory and confirmatory analyses, followed in the penultimate column by the loadings on the second order factor (identified as g).[2] The final column gives the loadings of the items obtained by Van

[2] The solutions shown in Table 2 are based on the weighted least squares exploratory factor analysis for categorical data as implemented in Mplus. In the confirmatory solution, all exploratory loadings <.20 were set at zero.

Table 2
Loadings of items on three-factor exploratory (weighted least squares) and confirmatory factors and on the higher order factor *g*

Items	Three-factor exploratory solution			Three-factor confirmatory solution			*g*	Van der Ven and Ellis (2000)
	F1: Gestalt continuation	F2: Verbal–analytic reasoning	F3: Visuospatial ability	F1: Gestalt continuation	F2: Verbal–analytic reasoning	F3: Visuospatial ability		
A1	**.57**	**.39**	−.23	** **	** **	** **		
A2	**.70**	.04	−.03	** **	** **	** **		
A3	**.69**	−.26	**.38**	** **	** **	** **		Gestalt
A4	**.73**	−.08	.10	** **	** **	** **		Gestalt
A5	**.76**	−.09	.08	** **	** **	** **		Gestalt
A6	**1.10**	.03	**−.45**	** **	** **	** **		Gestalt
A7	**.60**	.11	.23	**.68**	.00	−.14	**.68**	Gestalt
A8	.24	.06	.23	**.35**	.00	−.02	**.43**	Gestalt
A9	**.75**	.11	.19	**.57**	.00	.00	**.73**	Gestalt
A10	**.61**	.11	.08	**.30**	.00	.00	**.38**	Gestalt
A11	.26	.08	**.37**	**.35**	.00	.09	**.57**	Gestalt
A12	.16	.13	**.32**	** **	** **	** **		Gestalt
B1	**.78**	−.05	.05	** **	** **	** **		
B2	**.64**	.13	.20	**.84**	.00	**−.30**	**.67**	
B3	**.53**	−.14	**.55**	**.68**	.00	−.07	**.78**	Gestalt
B4	**.41**	−.04	**.45**	**.54**	.00	−.01	**.67**	Gestalt
B5	.06	−.02	**.59**	.00	.00	**.42**	**.56**	Gestalt
B6	.18	.01	**.35**	.00	.00	**.33**	**.44**	Gestalt
B7	.15	.02	**.34**	.00	.00	**.32**	**.43**	Gestalt
B8	.12	.04	**.48**	.00	.00	**.38**	**.51**	Analog
B9	.13	.02	**.54**	.00	.00	**.43**	**.58**	Analog
B10	.21	−.02	**.66**	.02	.00	**.50**	**.71**	Analog
B11	.17	.01	**.51**	.00	.00	**.42**	**.56**	Analog
B12	.06	.21	**.40**	.00	.15	.28	**.51**	
C1	**.37**	.03	**.49**	**.32**	.00	.18	**.72**	
C2	**.34**	.04	**.48**	**.35**	.00	.18	**.69**	
C3	**.33**	.05	**.42**	**.31**	.00	.16	**.62**	Analog+ LRDP
C4	.25	.11	**.32**	.24	.00	.16	**.53**	Analog+ LRDP
C5	.09	.19	**.50**	.00	.00	**.48**	**.65**	Analog+ LRDP
C6	−.03	.24	**.35**	.00	.14	.25	**.47**	Analog
C7	.17	.20	**.45**	.00	.09	**.43**	**.65**	
C8	.11	.16	**.37**	.00	.00	**.38**	**.51**	Analog+ LRDP
C9	.15	.09	.28	.00	.00	**.30**	**.40**	Analog+ LRDP
C10	−.06	**.37**	**.30**	.00	.23	.21	**.48**	Analog+ LRDP
C11	.04	**.41**	.25	.00	.27	.21	**.51**	
C12	−.01	**.59**	.06	.00	**.54**	.00	**.47**	Analog+ LRDP
D1	.14	−.35	**.78**	.00	−.51	**.78**	**.61**	
D2	−.06	−.01	**.77**	.00	.00	**.50**	**.68**	

Table 2 (*continued*)

Items	Three-factor exploratory solution			Three-factor confirmatory solution			*g*	Van der Ven and Ellis (2000)
	F1: Gestalt continuation	F2: Verbal–analytic reasoning	F3: Visuospatial ability	F1: Gestalt continuation	F2: Verbal–analytic reasoning	F3: Visuospatial ability		
D3	−.13	.05	**.79**	.00	.00	**.50**	**.67**	Analog
D4	−.03	.10	**.63**	.00	.00	**.49**	**.66**	Analog
D5	−.06	−.09	**.86**	.00	.00	**.52**	**.71**	Analog
D6	−.04	.03	**.74**	.00	.00	**.51**	**.68**	Analog
D7	−.19	.10	**.62**	.00	.00	**.39**	**.52**	Analog
D8	−.13	.19	**.56**	.00	.00	**.42**	**.56**	Analog
D9	−.13	.22	**.53**	.00	.12	**.33**	**.54**	Analog
D10	−.13	**.35**	**.57**	.00	.20	**.37**	**.68**	Analog
D11	.07	**.36**	.05	.00	.29	.00	.26	Analog
D12	.11	**.51**	−.05	.00	**.38**	.00	**.33**	Analog
E1	−.11	.26	**.48**	.00	.15	**.31**	**.55**	
E2	−.09	**.52**	**.36**	.00	**.30**	.27	**.63**	Analog
E3	−.07	**.49**	**.39**	.00	**.32**	.25	**.62**	Analog
E4	−.19	**.64**	.28	.00	**.38**	.17	**.57**	Analog+ C+UF
E5	−.17	**.71**	**.31**	.00	**.41**	.22	**.65**	Analog+ C+UF
E6	−.11	**.61**	.25	.00	**.38**	.17	**.56**	Analog+ C+UF
E7	.04	**.58**	.04	.00	**.42**	.00	**.37**	Analog
E8	−.03	**.64**	.12	.00	**.49**	.15	**.43**	Analog
E9	.05	**.72**	−.01	.00	**.50**	.00	**.44**	Analog
E10	.15	**.76**	−.16	.00	**.49**	.00	**.43**	Analog+ C+UF
E11	.15	**.80**	−.27	.00	**.52**	−.04	**.40**	Analog+ C+UF
E12	.02	**.72**	−.18	.00	**.42**	.00	**.37**	Analog
Expl. Var								

Factor labels according to van der Ven and Ellis: Analog—analogical reasoning; Gestalt—Gestalt continuation; LRDP—lack of resistance to perceptual distractors; C—coping factor; UF—unidentified factor; Expl.var—explained variance.

der Ven and Ellis (2000) on their gestalt continuation and analogical reasoning factors obtained from their Rasch analysis.

We now propose the identification of the three primary factors.

3.4.1. F1: Gestalt continuation

For this, the confirmatory solution should be regarded as definitive, but this can be supplemented with estimates from the exploratory solution. We identify the first factor obtained in both the exploratory and confirmatory factors as the cognition of figural relations factor of Gustaffson (1984, 1988) and the gestalt continuation factor of Van der Ven and Ellis (2000). While either term seems acceptable, we shall henceforth use gestalt continuation. This term is apt because the solutions to the items loading on the factor are most easily obtained by perception of the pattern as a gestalt and identifying the appropriate piece for its completion without the use of reasoning. In the present analysis, items A1 through A11 and

B1 through B4 are measures of gestalt continuation, while in the Van der Ven and Ellis study this factor is measured by items A3–12, B3–5, and more weakly by B6–7. In both data sets, responses to items A1 and A2 are highly skewed, which creates difficulties for analysis and classification. The explanation for this is that A1 is used for practice and the answer is explained and virtually everyone gets it right (in our data, 99.3% of participants gave correct answers). The same is probably true for A2 in the Van der Ven and Ellis data. The last item measuring gestalt continuation in the Van der Ven data is B7, while in our data the last item is B4. In our analyses, items C1, C2, and C3 also load on gestalt continuation although they do not load on this factor in the Van der Ven and Ellis analysis. Inspection of these items suggests that they can be solved by gestalt continuation, and we believe these results and their interpretation are straightforward.

One may argue that F1 is not interpretable as a gestalt continuation factor but an artefact of a ceiling effect that combines all items that are too easy to solve. In order to eliminate this interpretation, we found correlations between the mean percentages of correct answers of each item and their loadings on the three factors. The correlations were .44, .82, and .46, respectively, for gestalt F1, F2, and F3. Thus, the difficulty of items was the least important for F1.

3.4.2. F2: Verbal–analytic reasoning

We identify the second factor obtained in both the exploratory and confirmatory analyses as the verbal–analytic reasoning factor found by DeShon et al. (1995) in the APM. The justification for this identification is that the items loading on this factor are arithmetical addition and subtraction problems that require verbal reasoning for their solution. We believe that this is the only cognitive process that can be used to solve items C12, D12, E4, and E6–12 that have the highest loadings on this factor and are pure measures of it without any appreciable loadings on the other two factors.

3.4.3. F3: Visuospatial ability

We identify the third factor in both the exploratory and confirmatory solutions as the visuospatial ability factor found by DeShon et al. (1995) in the APM. The items with high loadings on this factor are B5 through B11; C5 and C7–9; D1 through D10; and E1. Inspection of these items suggests that solution can be found perceptually.

It is apparent that the analogical reasoning factor obtained by Van der Ven and Ellis (2000) has been split in the present analysis into two factors that we designate verbal–analytic reasoning and visuospatial ability. Items B12, C6, C10 and 11, D9–10, and E1–6 load appreciably on both these factors, indicating that both cognitive processes can be used to solve the problems.

As regards the second order factor, we follow standard practice in interpreting this as Spearman's *g*, the general factor on which all the items have positive loadings. Conventionally Spearman's *g* is equated with the first unrotated principal component (Jensen, 1998). In order to check our interpretation of the second order factor as *g*, we have calculated its correlation with the first principal component. The correlation is .99 and suggests that the two constructs are virtually identical, supporting the interpretation of the second order factor as *g*. However, the magnitude of loadings of the items on the second order factor does not show the usual pattern in which the more difficult items load more highly than the easy items on *g*. The most difficult items are E7–12, which have quite modest loadings in the range of .37 to .44 on the second order factor, as compared with a number of the easier items with loadings of .60 and above.

The correlations between the three primary factors are as follows: gestalt visualization and verbal–analytic reasoning correlate at .76; gestalt visualization and visuospatial ability at .86; and verbal–analytic reasoning and visuospatial ability at .71. The loadings of the three primary factors on the higher order factor are .95 (gestalt continuation), .80 (visuospatial ability), and .90 (verbal–analytic reasoning). Further, since interpretation of total scores on the Progressive Matrices is an issue of some importance because of the higher mean obtained by men (Lynn and Irwing, in press), we calculated the correlation between total scores and *g* and obtained a correlation .99.

We ran four factor exploratory and confirmatory solutions to see whether these might reveal the Van der Ven and Ellis' (2000) "lack of resistance to perceptual distractors" factor or the "coping" factor that they found in items E4–6 and E10–11. The four-factor solutions bring out the same gestalt continuation factor as appears in the three-factor solutions (items A1–11 and B1–5). There is also a verbal–analytic reasoning factor resembling the verbal–analytic reasoning factor in the three-factor solutions with high loadings of items C12, D11, and E8 –12. Factors three and four seem to be two further reasoning factors. Factor 3 has high loadings on items B8, B12, C6–8, C10, C11, D3, D8–10, and E1–6. Factor four has high loadings on B6, B7, B10, C1–4, D1, and D5. We are unable to offer any interpretation of Factors 3 and 4, but it is clear that they do not resemble the "lack of resistance to perceptual distractors" or the "coping" factors of Van der Ven and Ellis.

4. Age and sex differences

We now consider the sex differences on the three primary factors and on the higher order factor *g*. These are shown for three age groups with mean ages of 12.4 (n=768), 14.4 (n=744), and 16.8 (n=1223)

Table 3
Mean scores of boys and girls on (a) verbal reasoning (b) gestalt continuation (c) visual reasoning and (d) *g*, for three age groups

Age	Mean score		Standard deviation		Sample size		*d*
	Boys	Girls	Boys	Girls	Boys	Girls	
Verbal-analytic reasoning							
12 year olds	.447	.550	.229	.228	389	378	−.444***
14 year olds	.637	.652	.264	.261	374	370	−.057
17 years olds	.838	.780	.238	.234	508	715	.246***
Gestalt visualization							
12 year olds	1.089	1.120	.114	.107	389	378	−.192***
14 year olds	1.146	1.144	.072	.093	374	370	.024
17 years olds	1.189	1.176	.053	.075	508	715	.192***
Visuospatial ability							
12 year olds	1.168	1.245	.218	.190	389	378	−.370***
14 year olds	1.274	1.285	.146	.159	374	370	−.073
17 years olds	1.343	1.336	.010	.109	508	715	.072
g							
12 year olds	1.055	1.111	.151	.140	389	378	−.384***
14 year olds	1.147	1.151	.113	.127	374	370	−.033
17 years olds	1.221	1.203	0.85	0.99	508	715	.193***

*** P<.001.

Table 4
Analysis of variance for SPM factors by sex and grade

Source	*df*	MS	*F*	P
Verbal–analytic reasoning				
Sex (S)	1	0.245	4.22	.040
Grade (G)	2	23.237	399.33	.000
G×S	2	1.526	26.23	.000
Error	2728	0.058		
Gestalt visualization				
Sex (S)	1	0.018	2.50	.114
Grade (G)	2	1.444	198.11	.000
G×S	2	0.113	15.53	.000
Error	2728	0.0073		
Visuospatial ability				
Sex (S)	1	0.459	19.92	.000
Grade (G)	2	4.149	180.22	.000
G×S	2	0.425	18.47	.000
Error	2728	0.02		
g				
Sex (S)	1	0.108	7.81	.005
Grade (G)	2	5.22	379.51	.000
G×S	2	0.373	27.14	.000
Error	2728	0.014		

in Table 3. It will be seen that for all four factors, girls performed significantly better than boys at age 12, there is no significant sex difference at age 14, and boys performed significantly better than girls at age 17, with the exception of visuospatial reasoning, on which the male advantage was not significant. Table 4 shows by analysis of variance that there is a statistically significant Sex×Age interaction for all four factors.

5. Discussion

The results contain 11 points of interest. First, they confirm the position adopted by Gustaffson (1984, 1988) and Van der Ven and Ellis (2000) that the SPM is not a pure measure of reasoning ability and *g*, defined as the general factor or as reasoning ability; but that while most of the items are measures of reasoning, the early items are measures of a visualization factor that Gustaffson designated the cognition of figural relations and Van der Ven and Ellis have designated gestalt continuation. Although we have used different methods of statistical analysis from that of Van der Ven and Ellis, our results are closely similar in showing that the items in Set A and the initial items in Set B are entirely or mostly measures of gestalt continuation. Our results differ from those of Van der Ven and Ellis in detail in that our results identify items A1–10 and B2–4 as measures of gestalt continuation, whereas in the Van der Ven and Ellis analysis the items measuring gestalt continuation are items A3–12 and B3–7.

Second, our analysis differs from that of Van der Ven and Ellis (2000) in that it splits their analogical reasoning factor into two factors that we identify as verbal–analytic reasoning and visuospatial ability. These two factors are correlated at .71. We consider that this relatively modest correlation provides further justification for the splitting of the Van der Ven and Ellis (2000) reasoning factor into these two factors and corroborates the analyses showing that the three-factor model provides a much better fit than the alternatives.

Third, we were unable to replicate Van der Ven and Ellis's (2000) smaller factors that they identify as "lack of resistance to perceptual distractors" and "coping." Nothing resembling these factors appeared in our three-factor analysis or in our four-factor solution.

Fourth, the two factors we identify as verbal–analytic reasoning and visuospatial ability have been found in the APM by several investigators whose results are summarized in the introduction. We believe we have identified the same two factors in the SPM. Previous investigators who have found two factors in the APM have employed a number of different terms for these such as addition–subtraction and detection of pattern progression (Dillon et al., 1981) and reasoning and spatial ability (Colom & Garcia-Lopez, 2002). To prevent further proliferation of terms, we have adopted the names proposed by DeShon et al. (1995).

Fifth, it would be useful to identify the verbal–analytic reasoning and visuospatial ability factors found in this analysis of the SPM and by previous investigators in the APM with factors in the taxonomies of abilities of Carroll (1993) and McGrew and Flanagan (1998, p. 15). The verbal–analytic reasoning factor appears to be the same as fluid intelligence defined by Carroll (p. 626) and by McGrew and Flanagan (p. 15) as general sequential reasoning and defined by McGrew and Flanagan (p.15) as "the ability to start with stated rules, premises or conditions and engage in one or more steps to reach a solution to a problem." The visuospatial ability factor appears to be the visualization ability in Carroll's taxonomy, one of the Stratum 1 abilities in the second order broad visual perception factor. This factor is also present as a Stratum 1 ability in McGrew and Flanagan's (p.16) taxonomy, where it is a component of the second order visual processing factor and where it is defined as "the ability to rapidly perceive and manipulate visual patterns or to maintain orientation with respect to objects in space."

Sixth, the age by sex trends in our study show that for all four factors of gestalt continuation, verbal–analytic reasoning, visuospatial ability factors, and g, girls perform better than boys at the age of 12, there are no significant sex differences at the age of 14, while at the age of 17 boys perform better than girls, although the boys' advantage on visuospatial ability at the age of 17 is only 1.2 IQ points and is not statistically significant. This may be surprising because it is widely asserted that boys perform better than girls on tests of visualization and spatial abilities. However, Linn and Petersen (1985) in their meta-analysis of sex differences in these abilities proposed that there are three of these abilities that they identified as mental rotation, spatial perception, and spatial visualization. They concluded that males performed on average substantially better than females on mental rotation, somewhat better on spatial perception, but that there is little sex difference on spatial visualization. The factor we have identified as visuospatial ability in the SPM does not appear to involve mental rotation. It appears to be a measure of what Linn and Petersen called spatial visualization and hence our result that males have a nonsignificant advantage of 1.2 IQ points is consistent with their analysis.

Seventh, despite its factorial complexity, total scores on the SPM correlate so highly with g at .99 that they are indistinguishable from it. Thus, the contention that the Progressive Matrices is a relatively pure measure of g, somewhat paradoxically, appears to be true, at least in samples aged 12 or older. Among younger age groups, because gestalt continuation should have greater influence, *g* and total scores on the Progressive Matrices may show somewhat greater divergence. In the light of this finding, the male advantage in scores on the Progressive Matrices, which emerges at 15 years of age and increases to an average figure of .33 *d*s in adulthood (Lynn and Irwing, in press), may be equated to a male advantage in *g*.

Eight, our results confirm Jensen's (1998) contention that measures of reasoning and of *g* calculated as a higher order factor produce closely similar results. Thus, for our 17-year-olds, boys have an IQ advantage of 3.7 IQ points on verbal–analytic reasoning, while on *g* (calculated as a higher order factor) boys have an advantage of 2.9 IQ points.

Ninth, the sex by age trends found in this study provide further support for the developmental theory of sex differences in intelligence presented in Lynn (1994, 1999, 2002) and in Lynn, Allik, and Must (1999). This theory states that girls mature earlier than boys, and hence girls perform as well or better than boys at the ages of 9–13 years, while from the age of 16 boys perform better than girls. This shift from better performance by girls in early adolescence to better performance by boys in later adolescence is further testimony to the unsatisfactory procedure adopted by Court (1983), Jensen (1998), and Mackintosh (1996) of aggregating sex differences on the Progressive Matrices for all age groups, finding inconsistent results and concluding that there is no sex difference. The present results provide evidence that consistent sex differences are present when the data are analyzed by age.

Tenth, our results show that 17-year-old boys have an advantage of 3.7 IQ points in verbal–analytic reasoning and 2.9 IQ points in *g* calculated as a higher order factor. This is consistent with the contention first advanced in Lynn (1994) that among adults, males have an advantage of 3.8 IQ points in general intelligence calculated as the average of the verbal comprehension, reasoning, and spatial primary abilities. The male advantage among 17-year-olds is a little lower than among adults because the male advantage increases progressively from the age of 16 to adulthood and does not reach its maximum until adulthood. The research reported in Lynn (1994,1999, 2002) and by Lynn et al. (1999) shows that from the age of approximately 16 years, males have an advantage of 2.5–5 IQ points in general intelligence, whether this is defined as reasoning ability or as the average of the three major primary abilities of verbal comprehension, reasoning, and spatial abilities. The present analysis showing a male advantage of 2.9 IQ points in *g* calculated as a higher order factor provides further evidence for this conclusion. Other relatively recent results supporting the conclusion that among adults, males score a little higher on general intelligence, however this is defined, than females have been reported by Stumpf and Eliot (1995), Colom and Garcia-Lopez (2002), Lynn and Chan (2002), Nyborg (2003), and Deary et al. (2003).

Eleventh, Carroll (1993) and McGrew and Flanagan (1998) in their taxonomies of mental abilities list 11 first level visual/spatial abilities subsumed under the second order "broad perceptual processing" factor of Carroll and "visual processing" factor of McGrew and Flanagan. Neither of these taxonomies contain the first level cognition of figural relations factor identified by Gustaffson (1984, 1988) or the similar gestalt continuation identified by Van der Ven and Ellis (2000) and confirmed in the present analysis. We propose that this primary factor should be added to the 11 visualization and spatial abilities in the Carroll and McGrew and Flanagan taxonomies. According to our data, this ability develops relatively early and is more or less fully developed by the age of 12, unlike verbal–analytic reasoning and visuospatial ability, which continue to develop up to the age of 17.

References

Adcock, C. (1948). A re-analysis of Slater's spatial judgment research. *Occupational Psychology*, *22*, 213–216.

Alderton, D. L., & Larson, G. E. (1990). Dimensionality of Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement*, *50*, 887–900.

Arthur, W., & Woehr, D. J. (1993). A confirmatory factor analytic study examining the dimensionality of the Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement*, *53*, 471–497.

Banks, C. (1949). Factor analysis of assessments for army recruits. *British Journal of Psychology. Statistical Section*, *2*, 76–89.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement—A structural equation perspective. *Psychological Bulletin*, *110*, 305–314.

Bozdogan, H. (1987). Model selection and Akaike information criterion (AIC)—The general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 256–293). London: Sage.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika*, *40*, 5–32.

Colom, R., & Garcia-Lopez, O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality and Individual Differences*, *32*, 445–451.

Court, J. H. (1983). Sex differences in performance on Raven's progressive matrices: A review. *Alberta Journal of Educational Research*, *29*, 54–74.

Deary, I. J., Whalley, L. J., St.Clair, D., Breen, G., Leaper, S., Lemmon, H., Hayward, C., & Starr, J. M. (2003). The influence of the e4 allele of the apoliprotein E gene on childhood IQ, nonverbal reasoning in old age, and lifetime cognitive change. *Intelligence*, *31*, 85–92.

DeShon, R. P., Chan, D., & Weissbein, D. A. (1995). Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional performance determinants. *Intelligence*, *21*, 135–155.

Dillon, R. F., Pohlmann, J. T., & Lohman, D. F. (1981). A factor analysis of Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement*, *41*, 1295–1302.

Emmett, W. G. (1949). Evidence of a space factor at 11+ and earlier. *British Journal of Statistical Psychology*, *2*, 3–16.

Gabriel, K. R. (1954). The simplex structure of the progressive matrices test. *British Journal of Statistical Psychology*, *2*, 3–16.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, *68*, 532–560.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*, 265–275.

Gustaffson, J. (1984). A unifying factor model for the structure of intellectual abilities. *Intelligence*, *8*, 179–203.

Gustaffson, J. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence*, *vol. 4* (pp. 35–72). New Jersey: Laurence Erlbaum.

Hambledon, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer Academic Publishers Group.

Hertzog, C., & Carter, L. (1988). Sex differences in the structure of intelligence: A confirmatory factor analysis. *Intelligence*, *6*, 287–304.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance modeling: An overview and a meta-analysis. *Sociological Methods and Research*, *26*, 329–367.

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

Jensen, A. R. (1998). The *g* Factor. Westport, CT: Praeger.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). London: Sage.

Keir, G. (1949). The Progressive Matrices as applied to school children. *British Journal of Psychology. Statistical Section*, *2*, 140–150.

Kim, J. -O., & Mueller, C. W. (1978). *Factor analysis: Statistical methods and practical issues*. London: Sage.

Lim, T. K. (1994). Gender-related differences in intelligence: Application of confirmatory factor analysis. *Intelligence*, *19*, 179–192.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: A meta-analysis. *Child Development*, *56*, 1479–1498.

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences*, *17*, 257–271.

Lynn, R. (1999). Sex differences in intelligence and brain size: A developmental theory. *Intelligence*, *27*, 1–12.

Lynn, R. (2002). Sex differences on the Progressive Matrices among 15–16 year olds: Some data from South Africa. *Personality and Individual Differences*, *33*, 669–677.

Lynn, R., Allik, J., & Must, O. (1999). Sex differences in general intelligence among high school graduates: Some results from Estonia. *Personality and Individual Differences*, *26*, 1137–1142.

Lynn, R., & Chan, P. W. (2002). Sex differences on the Progressive Matrices: Some data from Hong Kong. *Journal of Biosocial Science*, *34*, 145–154.

Lynn, R., & Irwing, P. (2005). Sex differences on the Progressive Matrices: A meta-analysis. *Intelligence* (In Press).

Marsh, H. W., Balla, J. R., & Hau, K. -T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 315–335). Mahwah, NJ: Erlbaum.

Mackintosh, N. J. (1996). Sex differences and IQ. *Journal of Biosocial*, *28*, 559–572.

McGrew, K. S., & Flanagan, D. (1998). *The intelligence test desk reference*. Boston: Allyn and Bacon.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *43*, 551–560.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus. The comprehensive modelling program for applied research. User's guide*. Los Angeles: Muthén and Muthén.

Nyborg, H. (2003). Sex differences in *g*. In H. Nyborg, & J. Gray (Eds.), *The scientific study of human nature* (pp. 187–222). Amsterdam: Pergamon.

Panter, A. T., Swygert, K. A., Dahlstrom, W. G., & Tanaka, J. S. (1997). Factor analytic approaches to personality item-level data. *Journal of Personality Assessment*, *68*, 561–589.

Raven, J. (1939). *Progressive Matrices: A perceptual test of intelligence*. London: H.K.Lewis.

Raven, J., Raven, J. C., & Court, J. H. (2000). *Standard progressive matrices*. Oxford: Psychology Press.

Spearman, C., & Wynn-Jones, L. (1951). *Human ability*. London: Macmillan.

Spence, S. H. (1997). Structure of anxiety symptoms among children: A confirmatory factor-analytic study. *Journal of Abnormal Psychology*, *106*, 280–297.

Stumpf, H., & Eliot, J. (1995). Gender related differences in spatial ability and the *k* factor of general spatial ability in a population of academically talented students. *Personality and Individual Differences*, *19*, 33–45.

Van der Linden, W. J., & Hambledon, R. K. (1997). *Handbook of modern item response theory*. New York: Springer-Verlag.

Van der Ven, A.H.G.S., & Ellis, J. L. (2000). A Rasch analysis of Raven's standard progressive matrices. *Personality and Individual Differences*, *29*, 45–64.

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432–442.