Chapter 11

The Lithuanian Standardisation of the Coloured Progressive Matrices in an International Context*

Gražina Gintilienė, Dovilė Butkienė, and John Raven

Abstract

The *Coloured Progressive Matrices* (CPM) test, developed by J. C. Raven in 1947, is used world-wide to evaluate the non-verbal reasoning ability of 5 to 11 year-old children. This chapter presents Lithuanian CPM norms based on the standardisation conducted in 2004 of a representative sample of 6 to 11 year-old children (n=1067) in an international context: These norms differ, for example, from British and U.S. norms. An Item-Response-Theory-based item analysis, as well as the more usual split-half and internal consistency methods were used to assess the scaleability and validity of the test. It was found that environmental factors (place of residence of child and educational level of the parents) had a more significant effect in the development of non-verbal reasoning ability than the gender of the child. It is possible to conclude that both the test itself and the norms that were developed will be useful in evaluating the non-verbal reasoning ability of children for screening purposes.

*****

Raven's *Progressive Matrices* (RPM) are a series of measures which assess the ability to find meaning in confusing situations although they are often said to measure non-verbal reasoning ability. Already in 1930, while conducting research with mentally retarded subjects, J.C.

* The authors wish to thank Mrs Jean Raven for her help in generating of the smoothed percentiles.
We also thank Vida Gudauskiene, the psychologist of National Centre of Special Education and Psychology for her assistance in organisation of test administration at schools. Also we are grateful to psychologists of psychological pedagogical services for their voluntary work testing children and also to all Vilnius University masters degree educational psychology students participated in this study during the academic year 2004-2005.

Raven understood the need to develop a method which would allow the investigation not only of genetic factors affecting intelligence, but also environmental ones. That motivated him to develop a test whose results would be minimally dependent on acquired information but would be theoretically based and unambiguously interpretable. The initial version of the RPM appeared in 1938. Its theoretical base was Spearman's "***g***" factor which consists of two components: eductive and reproductive abilities. The term "eductive" comes from the Latin verb "educere" and means "the ability to make meaning out of confusion". The second term encompasses the ability to reproduce acquired information (Raven, 2000). Eductive ability includes the perception of a problematic situation and its analysis, the detection of problem, going beyond the given to perceive that which is not immediately obvious, forming constructs which make it easier to think about complex problems involving many mutually dependent variables. In the other words, it is the process of making sense of novel complex situations, when conclusions are based on active purposeful discovery and new insights rather than on simple choice among presented observable options. In solving the RPM the subject has to identify the missing detail of the picture, after determining the relationship among the elements presented in the matrix. The tasks become progressively more difficult. Reproductive ability encompasses the skill of learning, recalling, and reproducing verbal material. It is evaluated using Vocabulary scales, which are part of Raven's test battery (Raven, J., Raven, J.C., Court, 1998a).

The first version of RPM to be published was the *Standard Progressive Matrices* (SPM). Although the *Coloured Progressive Matrices* was developed at the same time, it was not published until 1947, at which time the *Advanced Progressive Matrices* (APM) was also published (although it had been developed for use by the armed services during the Second World War). The CPM comprises Sets A and B of the SPM with an additional set – Set Ab – interpolated between them. All CPM items are in colour. At this time, there are four versions of the RPM in use: CPM, SPM, APM and SPM*Plus* (SPM+). The SPM+ was developed to restore the discriminative power which the SPM had when it was first developed but which had been eroded by the so-called "Flynn effect". In 1998 parallel versions of SPM and CPM were developed to foil respondents who might have been coached in the correct answers (Raven et al., 1998a).

In developing the RPM, J. C. Raven tried to minimize the dependence of task solution on acquired knowledge, especially that acquired through

formal education. In research he carried out in 1936 he showed that the test worked – scaled – in the same way for children from all backgrounds. Nevertheless, there were major differences between the means and spread of the actual scores obtained by children from different socio-economic and educational backgrounds. A comparison of the 1979 and 1938 normative data for the *Standard Progressive Matrices* revealed that there had been a marked increase in the scores obtained by young people of the same age (Raven, 1981). Later, in 1987, Flynn reported that the population IQs of people from many countries had increased by 5 to 25 points from generation to generation. Moreover, the difference in mean IQ between generations was greater when non-verbal tests (such as the RPM) were used rather than when the assessments were made using multi-component tests such as the Wechsler or Stanford-Binet tests. That motivated researchers to look for other factors possibly influencing the results of the RPM. The latest research (at present the RPM have been standardised in more than 30 countries) does not provide a simple answer regarding factors influencing the development of reasoning abilities. In the opinion of Raven, J. (2000b) growth on the RPM score mean may be linked to the same factors as parallel increases in human height, birth weight, life expectancy, athletic ability, and the drop in mortality of infants - that is, to better nutrition, welfare, and hygiene. In addition, he noted that certain child-rearing practices and forms of education have a significant effect on the development of children's eductive ability. Those factors emphasise the importance not only of having separate standardisations for various countries, but also the need to re-evaluate the norms periodically because a country's intellectual potential changes with its economic and technical development. In fact, Aiken (2003) states that restandardisation is needed for all three RPM, including the CPM.

Until now not a single RPM version was standardised in Lithuania on a representative sample of the population. In 1999 259 children were evaluated in Kaunas, the second biggest town in Lithuania, using the CPM (Lynn & Kazlauskaite, 2002). Despite the fact that the study was not representative, being based on the results of children from some city schools only, it was concluded that the IQ of Lithuanian children is 94 and is lower than that of Russian children (IQ=97) and Estonian children (IQ=99). However, the first intelligence test, Wechsler Intelligence Scale for Children - third edition (WISC-III) was standardised in Lithuania only in 2002.

This unsatisfactory lack of standardised instruments prompted the current undertaking, which was intended to establish Lithuanian normative

data for, and examine the psychometric properties of, Raven's *Coloured Progressive Matrices* among 6 to 11 year-olds.

Methods

Participants

To select a representative sample of 6-11 year old Lithuanian children, several random sampling methods were used. First, by means of a stratified sampling method, centred specifically on such variables as place of residence (urban, town, or rural), native language (Lithuanian, Russian, or Polish), and the kind of educational setting (kindergarten or mainstream school), 79 educational institutions were selected. Further, out of each selected educational institution, by means of a simple random sampling method, 1-2 groups, or primary-school first classes (2nd, 3rd, and 4th ones) were selected. Finally, out of each selected class (group), based on its size, by means of a random sampling method, 2-6 children (boys and girls in equal proportion) fulfilling the age requirements for representative sample were selected. Demographic characteristics of representative sample (n=1067) are presented in Table 11.1. The distribution of children in total sample (also in each age group) by place of residence and gender closely corresponds with the data provided by the Lithuanian Statistics (Education, 2004).

The sample consists of 11 age groups, each of them covering one half-year. For example, the 6-year old group includes children whose age ranges from 5 years 9 months and 1 day to 6 years 2 months and 30 days. Similarly, the 6 ½ year olds' group includes children whose age ranges from 6 years 10 months and 1 day to 6 years 8 months and 30 days. Children of the same age, quite often, attend different grades, which is also true in case of 6- or 7-year olds among whom one may find preschoolers attending a kindergarten, and pupils attending 1st or 2nd grade at a primary school. Thus, the selection of children among selected class (group) pupils was based on the following criterion: The child had to be not younger than 5 years 9 months yet not older than 11 years 2 months. Table 11.2 presents the data on educational institutions attended by representative sample children, and on grades to which they belong.

Additional information on children's native language, educational programme, and parents' education was gathered from children's parents (caregivers) by means of a questionnaire. According to parents (caregivers) report, 87.7% children spoke only Lithuanian, 5.2% only

Table 11.1. *Coloured Progressive Matrices*
Demographic Characteristics of Lithuanian Sample

		Place of Residence						Gender			
		Urban		Town		Rural		Boys		Girls	
	n	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Age group											
6	50	21	42.0	16	32.0	13	26.0	25	50.0	25	50.0
6½	93	38	40.9	29	31.2	26	28.0	46	49.5	47	50.5
7	88	36	40.9	26	29.5	26	29.5	44	50.0	44	50.0
7½	98	40	40.8	29	29.6	29	29.6	49	50.0	49	50.0
8	88	36	40.9	26	29.5	26	29.5	44	50.0	44	50.0
8½	122	49	40.2	37	30.3	36	29.5	61	50.0	61	50.0
9	122	50	41.0	36	29.5	36	29.5	61	50.0	61	50.0
9½	115	47	40.9	34	29.6	34	29.6	58	50.4	57	49.6
10	95	39	41.1	28	29.5	28	29.5	48	50.5	47	49.5
10½	111	45	40.5	33	29.7	33	29.7	56	50.5	55	49.5
11	85	35	41.2	25	29.4	25	29.4	42	49.4	43	50.6
Gender											
Boys		218	40.8	160	29.9	156	29.3				
Girls		218	40.9	159	29.8	156	29.3				
Total	1067	436	40.9	319	29.9	312	29.2	534	50.0	533	50.0
Lithuanian population*			41.7		32.5		25.8				
Lithuanian population**			40.7		29.8		29.5				

* Percentage of 6-7 years old children attending preschool education groups at mainstream schools or kindergartens according their residential areas (Education, 2004).

** Percentage of children attending grades 1-4 in urban, town, and rural areas. (*Education*, 2004).

Russian, 3.7% only Polish, and 3.3% – several languages at home. Such distribution of children under study by native language corresponds with the data provided by the Lithuanian Statistics, 2004, on the number of Lithuanian children receiving education in different languages of instruction (Education, 2004).

Information on parents' education was reported by 934 (87.5% of the total number of children under study) parents (caregivers). Almost half the parents indicated they had post-secondary or higher education (44.6% fathers and 51.2% mothers). 5-6% of the parents' education was

Table 11.2. *Coloured Progressive Matrices*
Lithuanian Sample by Educational Setting

		Not attending kindergarten / school		Kindergarten		School							
						Grade 1		Grade 2		Grade 3		Grade 4	
	n	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Age group													
6	50	4	8.0	45	90.0	1	2.0						
6½	93	6	6.5	78	83.9	9	9.7						
7	88	5	5.7	34	38.6	48	54.5	1	1.1				
7½	98	2	2.0			86	87.8	10	10.2				
8	88					56	63.6	32	36.4				
8½	122					11	9.0	100	82.0	11	9.0		
9	122					2	1.6	80	65.6	40	32.8		
9½	115							10	8.7	101	87.8	4	3.5
10	95							2	2.1	71	74.7	22	23.2
10½	111									10	9.0	101	91.0
11	85					2	2.4			3	3.5	80	94.1
Total	1067	17	1.6	157	14.7	215	20.1	235	22.0	236	22.1	207	19.4

basic. The rest indicated that they graduated secondary or vocational school.

A representative sample of 6-11 year olds was selected including special needs children in mainstream schools. According to data provided by parents (caregivers), the majority of pupils attending 1st-4th grades were following a general education curriculum (96.1%), however the sample included pupils who studied according modified (2.4%), adapted (1.3%), or special individual (0.1%) programmes.

Description of the Coloured Progressive Matrices

The CPM consists of 36 coloured and attractive items: Three Sets (A, Ab, and B), each of 12 items, representing a series of patterns with a bit missing. Sets A and B correspond to the SPM Sets A and B, with Set Ab – drawn up for the CPM version – between them. Success in Set A is determined by individual's ability to complete a continuous pattern changing, first, in one direction, then, at the end of the set, in two directions at the same time. Success in Set Ab depends on individual's

ability to conceive discrete figures as a spatially related whole, and to choose a figure which completes the pattern. Set B is composed of items demanding ability to reason by analogy. Each of the 36 items requires that the child find a missing detail among six alternatives given at the foot of the picture, and to indicate it. The CPM is designed to assess the eductive ability of 5-11 year olds, mentally retarded people, and the elderly. By "eductive ability" is meant the ability to make meaning out of confusion; the ability to perceive. This involves simultaneously forging images of wholes and parts (this is not correctly rendered as "perceiving relationships between elements", because elements do not exist without wholes). It also involves the ability to reason by analogy. Some years after the first publication of the CPM, additional item analyses led to the modification of certain items and the order of presentation (Raven et al., 1998a).

Data on the reliability and validity of the CPM are available in the test Manual (Raven et al., 1998a). Split-half reliability estimates range from 0.65 to 0.94 (age groups from 6 to 8 years); retest reliabilities from 0.71 to 0.87 (5.7 and 8 year groups); and Cronbach alphas from 0.80 to 0.93 (5 and 11 ½ year age groups).

The validity of the CPM is supported by statistically significant correlations (from 0.50 to 0.80) with other "intelligence" tests (Terman-Merill, WISCR-R, Stanford-Binet) and achievement tests (Raven et al., 1998a), by scaleogram analyses and factor analyses of the correlations between numerous tests (e.g. Snow, Kyllonen and Marshalek, 1984, and Carroll, 1993) but, most importantly, by internal consistency analyses carried out on the test itself using Item Response Theory. As can be seen from other chapters in this volume, these show that the RPM tests have many of the properties of a "tape measure" or "meter stick" and that these are almost constant across many cultural groups.

Procedure

The standardisation was carried out in February – March and November – December, 2004. With the approval of Lithuanian Republic Ministry of Education and Science, the administration of educational institutions was informed about the study. Teachers were asked to hand a letter and questionnaire to children's parents (caregivers). The questionnaire was designed to obtain socio-demographic data (parents' education, child's native language, family composition), and other supplementary information (attending or non-attending a pre-school institution before

entering school proper, educational programme, additional education, etc.). Each child was tested with the prior written permission issued by his/her parents (caregivers).

Thirty eight specially trained primary researchers administered the booklet form of the CPM to selected children. The testing was conducted during morning hours at school in a room provided for the testing purpose by the manager of the educational institution. Kindergarteners and pupils attending 1st grade (like the children who did not attend any educational setting) were tested individually. Pupils attending 2nd grade were tested in pairs. Pupils attending 3rd and 4th grades were tested in groups including 4-6 children. The CPM administration took about 10 to 20 minutes.

As indicated in CPM Manual (Raven et al., 1998a), children unable to succeed on the first five items of Set A fail to realise how the test items are to be solved. As a result, their final score on CPM, whatever it may be, is considered to be invalid. So, the further data analysis was based only on results of those children who accomplished the first five tasks on the CPM.

Results

Norms

Means and standard deviations were generated at half a year intervals between ages 5 years and 9 months to 11 years 2 months. In order to eliminate sampling errors the raw scores for the main percentiles (5^{th}, 10^{th} etc.) for each age group were plotted and then smoothed by graphing across age groups. The Table 11.3 presents smoothed CPM norms for Lithuania.

There is a clear ceiling effect among older children (9 ½ +), and, even at the younger ages, the distributions are not symmetrical around the median score. For example, among 7 ½ year olds, the difference between the 5^{th} and 50^{th} percentile is 9 while that between the 50^{th} and 95^{th} is only 6. Because the test has only 36 items, the scores of the most able 10% of the population seem not to increase from 9 ½ years of age to 11 years of age.

It is obvious from these results that the test is unsuitable for use with more able children over nine years of age. As indicated in the Manual, those who wish to obtain reliable information on such children should (once their superior performance on the CPM has become apparent)

Table 11.3. *Coloured Progressive Matrices*
Smoothed Percentile Norms for Lithuania for 2004

	Age in Years (Months)										
	6	**6½**	**7**	**7½**	**8**	**8½**	**9**	**9½**	**10**	**10½**	**11**
	5(9)	6(3)	6(9)	7(3)	7(9)	8(3)	8(9)	9(3)	9(9)	10(3)	10(9)
	to	to	to	to	To	to	to	to	to	to	to
Percentile	6(2)	6(8)	7(2)	7(8)	8(2)	8(8)	9(2)	9(8)	10(2)	10(8)	11(2)
95	27	29	30	31	32	33	34	35	35	35	35
90	25	27	29	30	31	32	33	34	34	34	34
75	23	24	26	27	28	30	31	32	33	33	33
50	**21**	**22**	**23**	**25**	**26**	**27**	**28**	**29**	**30**	**30**	**30**
25	17	18	19	20	21	22	24	25	26	27	27
10	14	15	16	17	18	19	20	21	22	23	24
5	13	14	15	16	17	18	19	20	21	22	22
n	*50*	93	88	98	88	122	122	115	95	111	85

* Generated in collaboration with Jean Raven.

proceed to the *Standard* Progressive Matrices on completion of Set B and deduct the scores for Set Ab from the final scores obtained before seeking to convert them to percentile scores.

Group Differences

Table 11.4 compares the overall mean scores of boys and girls, assessing the statistical significance of the differences using the t test for independent samples. It is clear that there is no difference in the overall scores this group of young people aged 6 to 11 years.

An analysis of variance revealed significant differences by place of residence ($F(3, 1064) = 26{,}09$, $p < 0.001$, for raw scores and $F(3, 1064) = 36.26$, $p < 0.001$ for percentiles). As shown in Table 11.5 there are significant differences between children from urban areas and town or rural groups, with the children from the cities performing better than their peers attending town and village schools. The mean difference between last two groups is not statistically significant, although children from the towns do slightly better than rural children.

The analysis of the variation in scores with parental education was based on parents' answers to the questionnaire they had completed

Table 11.4. *Coloured Progressive Matrices*
Lithuanian Standardisation
Means and Standard Deviations by Gender

	Mean	Standard deviation	t	p
Raw scores				
Boys	26.34	5.63	1.151	0.250
Girls	25.94	5.64		
Percentile				
Boys	51.90	28.60	1.102	0.271
Girls	49.92	29.85		

Table 11.5. *Coloured Progressive Matrices*
Lithuanian Standardisation
Means and Standard Deviations by Place of Residence

	Mean	Standard deviation	t	p
Raw scores				
Urban	27.55	5.04	4.935	0.000
Town	25.57	5.93		
Urban	27.55	5.04	7.133	0.000
Rural	24.74	5.68		
Town	25.57	5.93	1.809	0.071
Rural	24.74	5.68		
Percentile				
Urban	59.38	26.96	5.479	0.000
Town	47.91	29.44		
Urban	59.38	26.96	8.257	0.000
Rural	42.15	28.97		
Town	47.91	29.44	2.478	0.013
Rural	42.15	28.97		

Table 11.6. *Coloured Progressive Matrices*
Lithuanian Standardisation
Means and Standard Deviations by Educational Level of Parents

Type of school (years of school completed)	Father					Mother				
		Raw Scores		Percentile			Raw Scores		Percentile	
	n	M	SD	M	SD	*n*	M	SD	M	SD
Basic (10 years)	46	22.98	5.67	33.70	25.44	57	22.68	5.82	31.98	25.82
Secondary (12 years)	170	24.86	5.83	42.74	28.41	211	24.82	5.60	42.21	28.11
Vocational (11-13 years)	265	24.88	5.53	45.09	29.12	180	24.85	5.64	46.93	29.45
Post-secondary (14-15 years)	210	26.90	5.28	55.34	27.06	257	26.30	5.55	51.67	28.81
Higher (college, university) (16-18 years)	177	28.07	5.25	66.27	26.62	224	28.04	5.17	63.99	26.52

Table 11.7. *Coloured Progressive Matrices*
Lithuanian Standardisation
Significance of Differences Between Mean Scores of Groups by Different Educational Level of Parents (see Table 11.6)

	Father's Educational level									
	B / S	B / V	B/ P	B / H	S / V	S / P	S / H	V / P	V / H	P / HI
Raw score										
t value	1.959	2.141	4.509	5.761	0.019	3.576	5.384	4.054	6.070	2.165
p value	0.051	0.033	0.000	0.000	0.985	0.000	0.000	0.000	0.000	0.031
Percentile										
t value	1.956	2.493	4.971	7.460	0.831	4.424	7.966	3.940	7.752	3.979
p value	0.052	0.013	0.000	0.000	0.406	0.000	0.000	0.000	0.000	0.000

	Mother's Educational level									
	B / S	B / V	B / P	B / H	S / V	S / P	S / H	V / P	V / H	P / H
Raw score										
t value	2.534	2.508	4.416	6.800	0.053	2.866	6.228	2.677	5.912	3.525
p value	0.012	0.013	0.000	0.000	0.958	0.004	0.000	0.008	0.000	0.000
Percentile										
t value	2.480	3.437	4.752	8.179	1.619	3.572	8.315	1.676	6.116	4.854
p value	0.014	0.001	0.000	0.000	0.106	0.000	0.000	0.094	0.000	0.000

B – Basic, S – Secondary, V- Vocational, P – Post-Secondary, H – Higher

about the school programme they had completed. ANOVA revealed significant effects for fathers ($F(4, 864) = 15.79$, $p < 0{,}001$, for raw scores and $F(4,864) = 25.84$, $p < 0.001$ for percentiles), and mothers ($F(4, 929) = 17.29$, $p < 0.001$, for raw scores and $F(4,929) = 24.58$, $p < 0.001$ for percentiles) education. Data presented in the Tables 11.6 and 11.7 show that the mean raw score of the pupils increases with parents' educational level. The differences between the groups are statistically significant. ($p<0.05$) except for groups of children whose fathers and mothers completed vocational and secondary schools.

Item Analyses

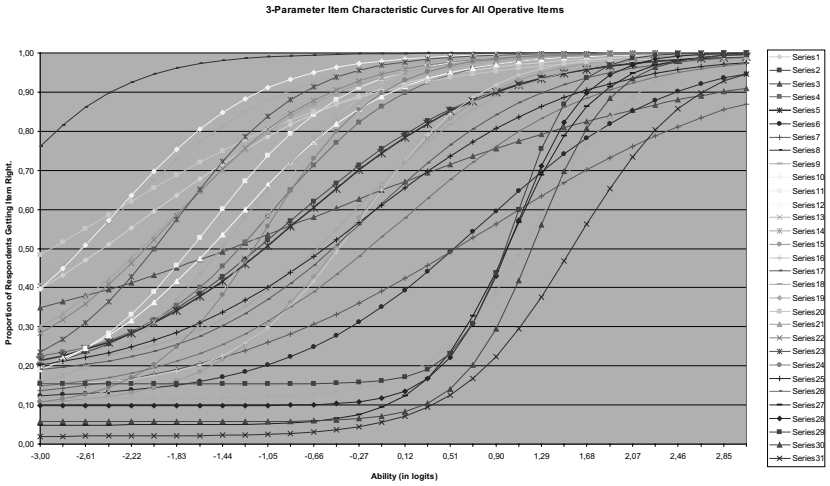
Split-half internal consistency correlations were calculated for the total sample and each separate age group and are presented in Table 11.8. All were high and ranged from 0.82 (seven year olds) to 0.90 (eight and half year olds and total sample). The internal consistency was also assessed using Cronbach's alpha (see Table 11.8). It can be seen that the Cronbach alpha coefficients are comparable to the split-half estimates: the lowest (0.76) was for six and seven year olds and the highest (0.84) was for the total sample.

As explained in the *General* Section of the RPM *Manual* (Raven et al., 1998, updated 2003) and in the *General Introduction* to this book, it does not make a great deal of sense to intercorrelate the items of tests developed according to Item Response Theory (IRT) and then use those correlations to calculate statistics like Cronbach's Alpha, still less as a basis for factor analyses in which an attempt is made to assess the "unidimensionality" of a test: Consider how meaningful it would be to conduct the same exercise based on the centimetre marks on a tape measure.

Table 11.8. *Coloured Progressive Matrices*
Lithuanian Standardisation
Split Half and Cronbach Alpha Coefficients for Total Sample and by Age Group

		Age Group										
	All	6	6½	7	7½	8	8½	9	9½	10	10½	11
Split Half	0.90	0.82	0.86	0.81	0.84	0.86	0.90	0.87	0.89	0.84	0.82	0.84
α	0.87	0.76	0.81	0.76	0.80	0.82	0.86	0.86	0.85	0.82	0.80	0.80

Figure 11.1. *Coloured Progressive Matrices*
Lithuanian Standardisation
3-Parameter Item Characteristic Curves for All Operative Items



Joerg Prieler was therefore commissioned to undertake a graphical 3-parameter Item-Response-Theory-based analysis along the lines discussed in other chapters of this book.

Figure 11.1 displays the Item Characteristic Curves for all the 31 items of the test that remained after the first five had been discarded because, as explained earlier, all respondents who got any of these items wrong were rejected from the analysis on the grounds that they had not understood what they were supposed to do.

It is immediately obvious that one item ... Number 3 in this plot, but which is actually A8 ... is seriously defective: Far too many low ability pupils get it right and far too many able pupils get it wrong. (It would not, however, be appropriate to modify this item in the test at this point in time since doing so would invalidate future comparisons with the vast amount of international and historical data that have been accumulated with the test as it stands.)

Also striking is the cluster of four items (three of which are Items 9, 10, and 11 from Set B) whose ICCs rise steeply toward the right hand side of the plot. What these tell us is that less able pupils hardly ever get these items right (except by "chance"), but, once pupils have developed the abilities needed to understand them, they seldom get them wrong.

Figure 11.2. *Coloured Progressive Matrices*
Lithuanian Standardisation
Test Characteristic Curve Calculated Over All Operative Items

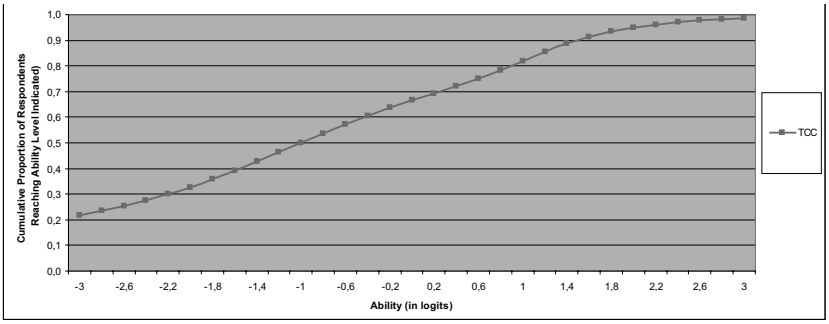
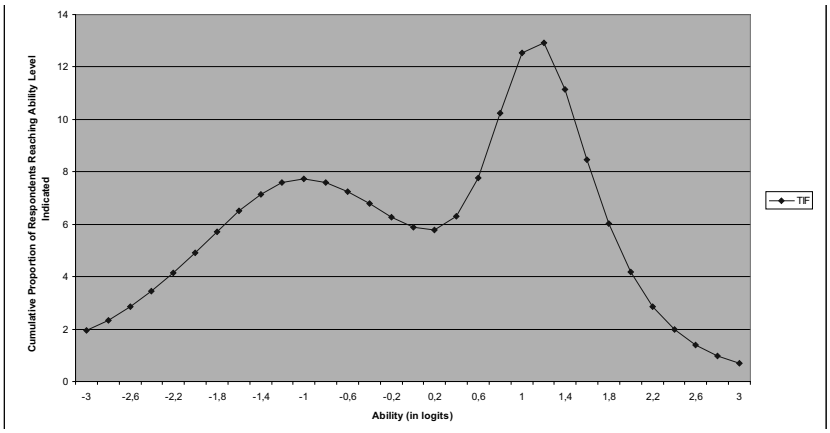


Figure 11.3. *Coloured Progressive Matrices*
Lithuanian Standardisation
Test Information Curve Over Operational Range of Test



It is also notable that that there are few items discriminating well in the middle ability range. The implications of this will become clear in a moment.

Figure 11.2 shows the Test Characteristic Curve for the Test as a whole. Whereas most people would expect it (and think it "ought") to approximate a Gaussian (often misleadingly termed "normal") ogive, it is surprisingly linear with few leaps or plateaux.

Figure 11.3 shows the Test Information Function curve.

At this point, a brief, if only approximate, explanation of what the Test Information Function (TIF) curve tells us may be offered (a suc-

cinct, but comprehensible and powerful, explanation will be found in Hambleton, et al., 1991). Basically, if there are many items with good discriminatory power (i.e. steep ICCs) at a particular level of ability covered by the test, the test yields a great deal of good information enabling users to accurately discriminate between those having that level of ability. If there are few items with good discrimination indices (i.e. having steep ICCs) at a particular point … usually at the top and bottom end of the range of ability for which the test is intended … little reliable diagnostic information can be gained from using the test. Thus if, as is often the case, users require a test which discriminates among those who appear to have particularly high or low ability, a test having the typical Gaussian-shaped Test Information Function curve will not give them what they want. A test having a rectilinear, or even bimodal, TIF would be of greater value. (In point of fact, of course, since the overall TIF does not reflect what happens within age groups, it would be desirable to calculate them separately within age groups. As it happens, however, as indicated above, a close inspection of percentile norms presented within age groups enables one to derive the relevant information directly.)

Be that as it may, the shape of the Test Information Function curve for the CPM shown above will disconcert many people. But, far from suggesting that the authors should tinker with the test to obtain a more Gaussian curve, the preceding discussion suggests that, if an attempt were to be made to improve the test, attention should concentrate on getting better discrimination in the upper and lower ability ranges.

Validity

Given normal cognitive development in children, one indication of the validity of the CPM would be a linear progression in raw scores over the age groups tested. The relationship between age and mean CPM raw scores was assessed by calculating a Pearson product-moment correlation coefficient. This was 0.52, thus accounting for 27.4% of the overall CPM total score distribution. As it is shown in Table 11.9 the mean of raw scores of 6-11 year old children increase with age (with the exception of ten and half and eleven year olds groups where, because of the ceiling effect noted above, more able children were unable to demonstrate their abilities).

Before the IRT based ICC analysis reported above had been commissioned, we assessed the content validity of the test on the basis of conventional item difficulty indices. Despite the more detailed information available from the plots of the ICCs, our analysis in terms of item difficulty

Table 11.9. *Coloured Progressive Matrices*
Lithuanian Standardisation
Means and Standard Deviations by Age Group

Total	Age Group										
	6	6½	7	7½	8	8½	9	9½	10	10½	11
Total											
M	19.94	21.29	23.23	23.61	24.78	26.08	27.77	28.09	28.85	30.01	29.41
SD	4.30	4.75	4.22	4.82	4.97	5.50	5.21	5.10	4.50	4.15	4.30
Set A											
M	8.34	8.88	9.22	9.32	9.44	9.61	10.04	10.07	10.36	10.46	10.38
SD	1.47	1.39	1.39	1.42	1.26	1.46	1.56	1.47	1.14	1.11	1.25
Set Ab											
M	6.58	7.28	8.11	8.09	8.58	9.23	9.63	9.75	10.08	10.46	10.34
SD	2.20	2.25	1.94	2.24	2.39	2.43	2.13	2.48	1.91	1.71	1.56
Set B											
M	5.02	5.13	5.90	6.20	6.76	7.25	8.10	8.27	8.41	9.09	8.69
SD	1.81	1.92	1.89	2.15	2.47	2.66	2.52	2.44	2.48	2.26	2.30
n	50	93	88	98	88	122	122	115	95	111	85

indices may nevertheless be of interest. The data presented in Table 11.10 reveal that there are some exceptions to a smooth progression in item difficulty. Items A9, Ab7, and B10 were slightly easier than one would expect and were more successfully solved than earlier items in correspondent Sets by children more than eight years old.

Discussion

In this paper we have reported the results of the first standardisation of the CPM on a Lithuanian representative sample. Internal consistency assessed by split half and Alpha coefficients are similar to those obtained elsewhere (Raven et al., 1998a).

However, as can be seen from Table 11.11, the Lithuanian norms do, however, differ from what has, in effect, become the international standard – the 1982 British norms - (Raven et al., 1998a). (They differ even more from the 1986 U.S. norms, which are low by international standards [Raven, 2000a]). If Lithuanian children were evaluated using the UK or US norms then, for example, the average 6 ½ year old would appear to score as well as the average British or American 8 year old. This does not necessarily show that Lithuanian children are more able than their British and American counterparts, but is probably due to the

Table 11.10. *Coloured Progressive Matrices*
Lithuanian Standardisation
Item Difficulties (p values) by Age Group

	Age Group											
Item	6-11	6	6½	7	7½	8	8½	9	9½	10	10½	11
A1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A6	0.96	0.94	0.94	0.97	0.96	0.94	0.93	0.94	0.97	1.00	1.00	0.99
A7	0.79	0.62	0.63	0.73	0.70	0.85	0.75	0.86	0.81	0.86	0.90	0.88
A8	0.76	0.56	0.69	0.78	0.74	0.68	0.71	0.75	0.77	0.90	0.84	0.81
A9	0.80	0.42	0.63	0.69	0.70	0.71	0.85	0.88	0.91	0.91	0.95	0.91
A10	0.77	0.54	0.65	0.72	0.68	0.68	0.74	0.82	0.86	0.86	0.88	0.89
A11	0.35	0.14	0.15	0.16	0.30	0.28	0.33	0.46	0.41	0.47	0.51	0.48
A12	0.30	0.12	0.19	0.17	0.24	0.30	0.30	0.34	0.35	0.36	0.39	0.41
Ab1	0.99	0.92	1.00	1.00	1.00	0.99	0.98	1.00	0.98	1.00	1.00	1.00
Ab2	0.97	0.94	0.96	0.96	0.95	0.98	0.97	0.98	0.97	0.95	0.97	1.00
Ab3	0.96	0.94	0.95	0.99	0.93	0.99	0.93	0.97	0.94	0.95	1.00	0.98
Ab4	0.87	0.74	0.71	0.86	0.79	0.82	0.84	0.92	0.90	0.95	0.96	0.94
Ab5	0.85	0.68	0.73	0.77	0.87	0.88	0.85	0.88	0.87	0.87	0.92	0.92
Ab6	0.76	0.44	0.46	0.60	0.71	0.77	0.78	0.82	0.81	0.88	0.92	0.94
Ab7	0.83	0.50	0.61	0.81	0.75	0.80	0.84	0.90	0.87	0.94	0.97	0.98
Ab8	0.63	0.28	0.33	0.39	0.46	0.55	0.73	0.77	0.72	0.82	0.80	0.75
Ab9	0.62	0.26	0.34	0.50	0.53	0.50	0.64	0.67	0.76	0.75	0.83	0.82
Ab10	0.63	0.38	0.44	0.51	0.45	0.56	0.70	0.68	0.74	0.78	0.78	0.72
Ab11	0.66	0.36	0.56	0.55	0.53	0.56	0.68	0.70	0.76	0.77	0.80	0.82
Ab12	0.32	0.14	0.18	0.18	0.13	0.21	0.29	0.35	0.44	0.43	0.50	0.47
B1	0.98	0.96	1.00	0.97	0.98	0.99	0.98	0.96	0.98	1.00	0.98	0.99
B2	0.93	0.86	0.83	0.91	0.89	0.93	0.92	0.98	0.94	0.97	1.00	0.94
B3	0.92	0.80	0.81	0.86	0.92	0.91	0.92	0.96	0.93	0.99	0.98	0.94
B4	0.90	0.86	0.76	0.85	0.87	0.91	0.89	0.95	0.91	0.94	0.94	0.95
B5	0.76	0.46	0.54	0.66	0.65	0.73	0.76	0.86	0.88	0.83	0.87	0.89
B6	0.64	0.32	0.51	0.56	0.53	0.56	0.61	0.75	0.71	0.73	0.80	0.69
B7	0.55	0.40	0.31	0.43	0.44	0.50	0.55	0.59	0.67	0.67	0.69	0.65
B8	0.33	0.04	0.05	0.09	0.21	0.24	0.30	0.41	0.46	0.47	0.55	0.54
B9	0.38	0.10	0.10	0.13	0.22	0.32	0.36	0.48	0.53	0.55	0.61	0.58
B10	0.46	0.12	0.14	0.26	0.26	0.32	0.47	0.56	0.63	0.58	0.73	0.67
B11	0.32	0.08	0.07	0.10	0.15	0.25	0.33	0.39	0.42	0.44	0.59	0.52
B12	0.17	0.02	0.02	0.08	0.08	0.11	0.16	0.20	0.21	0.24	0.33	0.33

internationally established intergenerational increase in scores (which has become known as the "Flynn effect) that has probably continued from 1982 into the present.

Although there was a tendency for boys to do better on the CPM than girls in the current study, no significant gender differences were found. This supports the British finding that the correlation of 0.01 between the SPM and sex is minimal (Raven et al., 1998b). In this context we may report that no significant differences were found between genders in the Performance IQ of the WISC-III, although the boys' Perceptual Organisation Index was significantly higher (Gintiliene & Girdzijauskiene, 2000). The results of the current investigation confirm previous findings that the same norms can be used in evaluating both boys and girls and that it is enough to maintain an equal number of males and females at various age groups in the standardisation sample.

In choosing the representative sample, a factor which was taken into account was place of residence. When the results of children living in different areas were examined, significant differences were found between those residing in large cities and those in rural areas. Similar results were found with the WISC-III (Gintiliene and Girdzijauskiene, 2003) when samples from rural and urban schools were matched on variables of gender, age and parents educational level. The Full scale IQ of rural children remained significantly lower. These findings may be related to income. The disposable income of rural population is 1.4 times smaller than that of residents living in large cities (Household Income and Expenses, 2004). These differences result in unequal opportunities for parents to develop children's abilities. In addition, the educational level of the parents is another socio-economics factor influencing the results. The WISC-III standardisation results revealed a statistically significant correlation ($r=0.32$, $p<0.01$) between a child's IQ and the educational level of his parents (Gintiliene and Girdzijauskiene, 2003). This relationship also has been pointed out by Sattler (2001) and others. The current research confirmed the fact that the lower the parents' educational level was, the lower were the child's CPM results. No difference between CPM results of children whose parents completed secondary school and those whose parents finished vocational school was found. This may be due to similarity of curricula between them. This fact also confirms CPM results' dependence on parents' educational level.

Table 11.11. *Coloured Progressive Matrices*
2005 Lithuanian Norms in the Context of 1982 British and 1986 American Norms

	Age in Years (Months)											
		6			6½			7			7½	
	5(9) to 6(2)	5(9) to 6(2)	5(9) to 6(2)	6(3) to 6(8)	6(3) to 6(8)	6(3) to 6(8)	6(9) to 7(2)	6(9) to 7(2)	6(9) to 7(2)	7(3) to 7(8)	7(3) to 7(8)	7(3) to 7(8)
Percentile	UK	US	**LT**	UK	US	**LT**	UK	US	**LT**	UK	US	**LT**
95	24	25	**27**	26	28	**29**	28	30	**30**	31	31	**31**
90	21	23	**25**	23	25	**27**	25	27	**29**	28	29	**30**
75	19	19	**23**	20	21	**24**	21	23	**26**	23	25	**27**
50	**16**	**14**	**21**	**17**	**16**	**22**	**18**	**18**	**23**	**20**	**20**	**25**
25	13	12	**17**	14	13	**18**	16	14	**19**	17	15	**20**
10	11	10	**14**	12	11	**15**	13	12	**16**	14	13	**17**
5	9	9	**13**	11	9	**14**	12	10	**15**	13	11	**16**
n	23		**50**	42		**93**	54		**88**	55		**98**

	Age in Years (Months)											
		8			8½			9			9½	
	7(9) to 8(2)	7(9) to 8(2)	7(9) to 8(2)	8(3) to 8(8)	8(3) to 8(8)	8(3) to 8(8)	8(9) to 9(2)	8(9) to 9(2)	8(9) to 9(2)	9(3) to 9(8)	9(3) to 9(8)	9(3) to 9(8)
Percentile	UK	US	**LT**	UK	US	**LT**	UK	US	**LT**	UK	US	**LT**
95	32	32	**32**	33	33	**33**	34	34	**34**	35	35	**35**
90	30	30	**31**	32	31	**32**	33	32	**33**	33	33	**34**
75	25	27	**28**	27	29	**30**	29	30	**31**	31	31	**32**
50	**22**	**22**	**26**	**24**	**24**	**27**	**26**	**26**	**28**	**28**	**27**	**29**
25	18	17	**21**	20	19	**22**	22	21	**24**	24	22	**25**
10	15	14	**18**	16	15	**19**	17	16	**20**	19	17	**21**
5	14	12	**17**	14	12	**18**	15	13	**19**	16	14	**20**
n	44		**88**	48		**122**	52		**122**	37		**115**

(continued)

Table 11.11. *Coloured Progressive Matrices*
2005 Lithuanian Norms in the Context of 1982 British and 1986 American Norms *(continued)*

	Age in Years (Months)								
	10			10½			11		
	9(9) to 10(2)	9(9) to 10(2)	9(9) to 10(2)	10(3) to 10(8)	10(3) to 10(8)	10(3) to 10(8)	10(9) to 11(2)	10(9) to 11(2)	10(9) to 11(2)
Percentile	UK	US	**LT**	UK	US	**LT**	UK	US	**LT**
95	35	35	**35**	35	35	**35**	35	35	**35**
90	33	33	**34**	34	34	**34**	35	34	**34**
75	32	32	**33**	33	32	**33**	33	33	**33**
50	**30**	**28**	**30**	**31**	**29**	**30**	**31**	**30**	**30**
25	25	23	**26**	26	24	**27**	28	25	**27**
10	21	18	**22**	22	19	**23**	23	20	**24**
5	17	15	**21**	18	16	**22**	20	17	**22**
n	53		**95**	49		**111**	51		**85**

Conclusions

Lithuanian norms for the Raven's *Coloured Progressive Matrices* were developed using data collected from on a representative sample 6 to 11 year-olds. These norms differ from their British and U.S. equivalents.

The study reconfirmed the reliability and validity of the CPM. The norms that have been established can therefore be used with confidence to evaluate the non-verbal reasoning ability of Lithuanian children in the course of formal assessments or screenings. Children living in rural areas and having less educated parents in general scored less well on the CPM. Gender differences were not significant.

References

Aiken, L. R. (2003). *Psychological Testing and Assessment*. Boston: Pearson Education Group.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. New York: Cambridge University Press.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measures. *Psychological Bulletin, 101*, 71-191.

Gintiliene, G., & Girdzijauskiene, S. (2000). Gender difference in intelligence of Lithuanian children. 11th European Conference on Personality. Freidrich-Schiller – Universitat Jena, 2000, July. *Conference Program and Abstracts* (p.105). *Lengerich: Pabst. Science Publishers*.

Gintiliene, G., & Girdzijauskiene, S. (2003). Lithuania: Culture and children's intelligence. In J. Georgas, L.G. Weiss, J. R. Van de Vijver, & D. H. Saklofske (Eds.), *Cross-cultural Analysis of the WISC-III* (pp.165-179). San Diego: Academic Press.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Press.

Lynn R., & Kazlauskaite V. (2002). The study of IQ in Lithuania. *Perceptual and Motor Skills, 95(2),* 6111- 6112.

Neisser, U. (Ed.). (1998). The Rising Curve. Washington, DC: American Psychological Association.

Raven, J. (1981). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No. 1: The 1979 British Standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales, Together With Comparative Data From Earlier Studies in the UK, US, Canada, Germany and Ireland*. San Antonio, TX: Harcourt Assessment.

Raven, J. (2000a). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No. 3 (Second Edition): A Compendium of International and North American Normative and Validity Studies Together with a Review of the Use of the RPM in Neuropsychological Assessment by Court, Drebing, & Hughes*. San Antonio, TX: Harcourt Assessment.

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

Raven. J., Raven J. C., Court J. H. (1998a). *Manual for Raven's Progressive Matrices and Vocabuliary Scales, Section 2: Coloured Progressive Matrices*. San Antonio, TX: Harcourt Assessment.

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

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence*, Volume 2, (pp. 47-103). Hillsdale, NJ: Lawrence Erlbaum Associates.