Raven's Standard Progressive Matrices: Contribution to Italian standardization for subjects between ages 6 and 18

Laura Picone, Arturo Orsini, Lina Pezzuti

Department of Dynamic and Clinical Psychology, Sapienza University of Rome

ABSTRACT. Il lavoro presenta la taratura delle *Standard Progressive Matrices* di Raven su un campione italiano di 5438 ragazzi dai 6 ai 18 anni. Lo strumento è molto conosciuto e datato ma tuttora utile per una rapida valutazione delle abilità cognitive, confermata nella letteratura internazionale dal confronto con i risultati di altri test che misurano lo stesso costrutto. Il presente lavoro di taratura ha confermato l'aumento delle prestazioni dei ragazzi al test con il crescere dell'età ma anche l'influenza sulle stesse di variabili sociali, come la scolarità dei genitori presa come indicatore del livello socio-economico familiare. I punteggi grezzi ottenuti dal campione sono stati trasformati in punti standard per età (QI con M = 100 e DS = 15); sono stati calcolati l'attendibilità e l'errore standard di misura; tali parametri sono riportati nel Report di Giunti Testing e servono a dare una corretta interpretazione dei punteggi ottenuti.

SUMMARY. This paper presents a standardization study of Raven's Standard Progressive Matrices on an Italian sample of 5438 young people aged 6 to 18 years. This instrument is well known, and although it is dated it is still useful for fast assessment of cognitive abilities, as confirmed in international literature a comparison of results from other tests which measure the same construct. The study confirms that young people performance is related with age, and it is influenced by social variables, such as parents' schooling level used as an indicator of social-economic level. The raw scores obtained by the sample were converted in standardized points for age (IQ with M = 100 and SD = 15). Reliability and standard measurement error were calculated. These parameters are included in the report by Giunti Testing; their purpose is to supply a correct interpretation of the scores obtained.

Keywords: *Intelligence, Raven's Standard Progressive Matrices, Youth, Adolescence, Cognitive Function*

INTRODUCTION

The first intelligence tests were created and spreaded at the beginning of the 1900s (Binet-Simon Scale, Binet and Simon, 1908; Stanford-Binet Intelligence Scales, Terman, 1916). They were criticized, however, because, being based primarily on verbal tests, they tended to penalize subjects belonging to culturally disadvantaged contexts. The first nonverbal tests were created by Otis (1936-1939), Wechsler (1939), and Raven (1938). In particular, the first version of Raven's Progressive Matrices (1938) was highly appreciated, coming at a time when the effort was to capture the essence of intelligence as independent of the knowledge which an individual has from schooling and from interaction with the surrounding environment.

John Carlyle Raven (1902-1970) had an idea for how to probe the question of intelligence without relying on verbal stimuli. He had been a student of Spearman's, the author of an intelligence theory (1927) which hypothesized the presence of a "general" and a "specific" factor. Together with the geneticist Penrose, Raven used Spearman's stimuli again; they did not, however, ask subjects to identify solutions verbally, but rather to find visually the solution which would fill in the incomplete abstract figures presented. Specifically, subjects had to identify a recurring pattern among the figure-stimuli based on an "inductive reasoning" method which apparently did not depend on scholastic or environmental education or knowledge.

Raven's Matrices are considered to be the best measure of Spearman's *g* factor, on a genetic and hereditary basis (Kaplan & Saccuzzo, 1997; Raven, 1938; Raven, Raven & Court, 1998).

The first version of the test, the *Standard Progressive Matrices* (*SPM*), was devised by Raven to measure cognitive capacity in children, adolescents, adults, and the elderly, even though national and international standardization data for this last category is more scanty.

Raven standardized them on a sample of Scottish young people (1407 subjects) aged 6 to 14 (1938, 1940). In 1944, Raven and Walshaw published another work on a sample with equivalent ages, from Colchester, above all to test correlations between the SPM and a vocabulary test (*Mill Hill Vocabulary Scale - MHV*). They found that the norms were 2 points lower than those from 1938. Raven attributed this datum to the restrictions and limitations the war-time period imposed, and this lead him to the conclusion that the test results were sensitive to environmental conditions and changes, both in terms of nutritional deprivation and in terms of a reduction in significant stimuli to which the young people were exposed.

Raven's son, John Raven Junior, got involved with the studies and carried them forward after his father's death. In 1972, with his two brothers, he established a company to manage the Matrices' dissemination in countries around the world.

Raven's SPM have been highly successful throughout the world, above all in English-speaking countries. Many researchers have done studies to verify whether the British norms can be used as a reference standard. Here below we cite some of the most important ones. Adams (1952) records the norms of 11621 young people aged 12 from Surrey which, with the limits of measurement error, are very similar to those gathered for the Scottish young people in 1938. Tuddenham, Davis, Davison, and Schindler (1958) tested various classes of children in California, and concluded that the use of these normative data is acceptable. Byrt and Gill (1973), working with Raven, gathered a representative sample of children aged 5-11 (3464) in the Republic of Ireland; the results for city children were comparable to the 1938 standards, even if those for rural-based children were slightly lower. Kratzmeier and Horn (1979) reported standards from a broad study done in Germany, with results well above those obtained in England in 1938. As mentioned, this lead John Raven Junior to develop a new British standardization (1979, published in 1981) on a sample of 3569 youth aged 6 to 16, where greater attention was paid to the representativeness of the sample.

In the United States 50 standards studies have been done (between 1983 and 1989, reported in J. Raven et al, 1990/2000) on a population drawn from each school district. Approximately 60000 students aged 5 to 18 were tested. It was shown that American standards for the white population are similar to the British standardization, but some ethnic groups obtain lower average scores.

Other authors report similar results to the English norms of 1981: Holmes (1980) in British Columbia (Canada); de Lemos in New Zealand and in Australia (1984, 1989); Abdel-Khalek and Raven in Kuwait (2006). Zhang and Wang (1989) in China, gathering data on an urban and rural population, found similarities with the Chan standards (1981, 1989) for Hong Kong, which corresponded closely with those obtained elsewhere. The same result is to be found in Poland (Jaworowska & Szustrowa, 1991), in Spain (Raven, Court & Raven, 1995); and in Switzerland (Martinolli, 1990).

To conclude, research appears to confirm that the

reference standards found in other countries are similar to those from the last British standardization. Furthermore, studies also seem to confirm that test results are connected to social-cultural characteristics of the social group in which the subjects live, given that results from populations in privileged environments generally seem better than those from less-privileged ones. This point revives the old debate regarding the weight and importance of hereditariness or environment on the development of mental abilities. It appears that improvement (or worsening) of a population's living conditions may induce a positive (or negative) change in the development of its members' intellectual abilities.

In recent works, Raven Jr. (2000, 2008) discusses the stability and the variations in SPM standards among cultural, ethnic, and social-economic groups over the last 70 years, which were also revealed by other researchers. The comparison between the new British standards and the original ones highlights a score increases over time for the Standard Progressive Matrices.

Flynn's studies raised a lot of interest in this regard. An early study of his (1984), starting with an analysis of SPM results obtained from various cohorts of military recruits, revealed a consistent increase in test scores over the years. Later, in 1987, he compared results on a number of intelligence tests (in addition to the Standard Progressive Matrices, also the Wechsler scales) from 14 countries in Europe and America (including Raven's studies on 30000 students aged 5 to 18) with his studies did in Australia on a population of various ages, including children and adults. Flynn found that score increases over time were to be found above all in standards studies of the Standard Progressive Matrices and not in other tests, quantifying this increase to be from 5 to 25 IQ points every 10 years, with differences among the various countries. In any case, the extent of the increase varies among the studies of various authors. Some of them showed an increase of almost 7 points per decade in SPM scores (e.g. Flynn, 1998). This increase, which has since been called "the Flynn Effect", is thought to be due to environmental factors such as improved nutrition, living conditions, and stimuli among which could be included the advent and use of technologies which broaden the possibilities available to individuals for accessing and actively seeking out knowledge and information (such as the television and computer).

In Italy, too, the SPM are rather well-spread, but with the use of standardizations on smaller samples which have not always been representative of the Italian population. Ferracuti and Groppelli (1954) and Boschi (1960) tested youth attending middle school (596 subjects between ages 11 and 13) and vocational-technical school (263 subjects between ages 12 and 14). The authors obtained the same results. Whereas gender showed no incidence, they did find significant differences for age and for the two types of school, with higher results among the youth from middle school who generally also belonged to a higher social-cultural class. They concluded, therefore, that the test's result is influenced by cultural variables. Other Italian studies highlight the effect one's level of schooling has on test performance. For example, Di Fiore and Renda (1968) (with a sample of 1560 youth aged 14 to 20) and Reda, Nencini and Riccio (1955) (with a sample of 400 subjects aged 16 to 29), observed higher performance among university graduates.

An Italian standardization study was published by Valseschini and Del Ton in 1973 on a sample of 1123 (990 males and 133 females) with ages ranging from 11 to 60, of which 857 subjects were between 11 and 20 years of age; their level of schooling went from illiterate to high school studies. In general the authors underline the absence of either a gender or age effect (the latter ostensibly being due to the sampling), and the presence of a schooling-level effect on the SPM. Starting with the group aged 31-40 years, a decrease in elderly people's performance is seen, which is mitigated by effect of schooling. Nevertheless, it should be noted that their Italian standardization sample, while being quite numerous, covered a broad age range, and the numbers in the various age ranges were not numerically balanced. Furthermore, subject distribution for gender was uneven, favoring the males.

One more recent Italian study (Giunti O.S., 2008) supplies reference standards for 825 non-clinical subjects aged 11 to 14, who go to middle or high school (defined as a non-clinical group) and 170 subjects aged between 10 and 84 years, defined as a clinical group given that they have various pathologies (e.g. trisomy 21, Alzheimer's). A standardization was calculated through division into percentiles for each of the two groups, leaving aside the fact that each clinical subject should always be compared to the standard. The clinical sample could be used only comparing it with the non-clinical sample and supplying the average performance of the two samples, after having verified that the two samples were paired at least for the most important variables like age, gender, and schooling.

The main results in the non-clinical group (11-14 years) are: an increase in scores compared with the 1973

standardidation, just as was seen at the international level; the perception that the SPM are easy, based on the difficulty index calculated on the test's 5 series (naturally, the items which are perceived to be difficult are the last ones in each series and the last series); the discrimination indexes for satisfactory items; the calculated reliability looking at internal consistency (equal to .90 with the Kuder-Richardson coefficient calculator); there are no significant differences for gender except in series D.

The validity of Raven's matrices as an instrument for the measurement of mental abilities has been studied through comparison with the results of other tests with the same construct and has been confirmed by a substantial number of studies. The first data were summarized by Burke (1958) and then reported by Raven, Court and Raven (1983; 1986; 1992), and by Raven and Summers (1986). Burke found that the correlation between the SPM and the Wechsler scales (which provide a mental efficiency index, FSIQ) grows among elderly subjects, while the validity coefficient with development tests for children varies from .30 to .60. In the Anglo-Saxon studies, these values are lower than those found when correlating the SPM with the Wechsler and the Binet Scales, administered immediately before or after the SPM (they vary from coefficients of $r = .54$ to $r = .86$; Raven, Court and Raven, 1978). The correlation of the SPM and these instruments tends to be higher with the performance subtests than with the verbal ones. In a large study involving thousands of students, Saccuzzo and Johnson (1995) conclude that the SPM and the WISC-R have approximately the same predictive validity in regard to criteria such as school performance, and that there is no different validities among 8 different ethnic groups were found.

Factor analysis studies concerning the SPM have highlighted the presence of a *general* factor, the relevance of which is more or less important depending on the study, and a saturation in s (the ability to visualize spatial relationships) has also been found.

Van der Ven and Ellis (2000), in a study aimed at identifying talented youth, administering the SPM form, used the Rasch model to investigate the items' one-dimensionality. In items in the A, C, and D series one-dimensionality was confirmed, whereas in the B and E series the items reveal two different dimensions.

With a factor analysis study (on 2735 youg people aged 12 to 18 in Estonia) Lynn, Allike, and Irwing (2004) explore and confirm the presence of a secondary factor g, but they also come across the presence of 3 other factors: Gestalt continuation (which had already been found by Van der Ven & Ellis, 2000), verbal-analytical reasoning, and visual-spatial ability.

Also Mackintosh and Bennett (2005) find (among 97 college students aged 17/18 years) that the g-factor is mainly present in the SPM, but the easiest items measure a perception or Gestalt factor which is distinct from an analytic factor in the rest of the test.

In an Italian study by Picone (1996), with 2000 subjects aged 14 to 19 years, factor analysis on concrete and formal Piagetian tests and on the SPM shows that the latter are saturated for the factor which includes the formal tests (.54) and for those including concrete tests (.44), as further proof for the fact that the kind of reasoning found in the Matrices can be considered a general intelligence factor.

The aim of this paper is to describe the fundamental data which were used for a new Italian standardization of the SPM using a sample of 5438 subjects ranging from 6 to 18 years of age. Some validation studies done on the sample are also presented.

METHOD

Sample

The Italian standardization sample is made up of 5438 Italian subjects (2410 males and 3028 females) aged 6 to 18 years, devoid of any evident psycho-physical handicap. Table 1 reports subject distribution according to 13 age levels, of one year each (e.g. from 6 years 0 months and 0 days to 6 years 11 months and 30 days), and according to gender.

In addition, subjects were classified in three groups according to their father's years of schooling: 2538 subjects had fathers who did up to 8 years of schooling; 2235 subjects' fathers did between 9 and 13 years; and 665 subjects' fathers did more than 13 years.

For all subjects, Raven's Standard Progressive Matrices were administered collectively and with no time limit. For the younger children, administration was done in small groups.

Subjects for the sample were recruited in various regions of Italy as part of work on university theses on the intelligence efficiency of children and younger people. Subject groups were given other tests of similar construct, along with the SPM: Human figure drawing for cognitive assessment,

Table 1 – Study sample configuration by age and gender

	Gender		
Age	**F**	**M**	**Total**
6.0-6.11	141	138	279
7.0-7.11	175	163	338
8.0-8.11	224	188	412
9.0-9.11	186	220	406
10.0-10.11	163	190	353
11.0-11.11	145	141	286
12.0-12.11	84	74	158
13.0-13.11	197	173	370
14.0-14.11	285	208	493
15.0-15.11	676	438	1114
16.0-16.11	246	181	427
17.0-17.11	206	134	340
18.0-18.11	300	162	462
6.0-18.11	3028	2410	5438

Piagetian tests for concrete and formal thinking, aptitude tests for school orientation.

The area from which data was collected can be defined as central-southern Italy: approximately half of the subjects came from Lazio (47%), but Campania (13%) and Sardinia (15%) are also well-represented.

Instrument

– *Description of the SPM*. The Standard Progressive Matrices (SPM) were the first form developed by Raven, in 1938, and for this reason are also known as PM38.

The materials needed for administration of the test are made up of a test booklet, an answer sheet with a grid on it, and a pencil (so as to avoid having permanent marks which could reveal correct answers in the booklet).

The exercise which subjects take on concerns solving logic problems, using graphical stimuli which are organized in matrices (2 x 2 or 3 x 3) made up of 60 items; these are divided in five series (A, B, C, D, E), increasing in difficulty within each series and from one series to the next.

The subject is asked to identify the figure which completes the item from among the alternatives. To do this, the subject has to observe the characteristics of the figures in the matrix both horizontally and vertically, as well as their relationship one to the other, given that the figures include graphical patterns which change from left to right, following some specific rationale, and from top to bottom, following some other rationale.

In general, the test requires that a series of concepts found in a set of problems of visual analogies be analyzed, constructed, and integrated with one another. Each entry (cell) contains one of the figurative elements, various kinds of shapes, various kinds of lines, or background patterns.

Incorrect answers come from an unclear perception of the constant elements and the modified ones in the test figures. The younger children tend to look for equivalencies

and identify the correct answer in alternatives which have elements that are the same as those in the stimulus. Children aged 8 to 10 can keep only one peculiarity of the figure in mind, but not both at the same time. Older children and youth (between ages 11 and 18) can consider both of the abstract figure's characteristics which change, and so get to the correct answer. Examining these observations in light of Piaget's theory of cognitive development, one can identify a typical mental process of the *preoperational* stage in the first group; the period for the *concrete operational* stage in the second group; and the presence of logical processes belonging to the *formal operational* stage in the third group. Using these reference parameters, an in-depth qualitative analysis could be done on an individual's performance.

– *SPM Administration Process*. The SPM can be administered collectively or individually, with no time limit. For younger children and for the elderly, individual or small-group administration is recommended, so as to check that basic instructions are followed well, both for the task at hand and for the response procedures.

The examiner, in the case of collective testing, will give the following instructions:

"Open your booklet to the first page: you will find a figure like this one"; at the same time a booklet open to the first page is shown to the whole group, while adding:

"At the top, on this page, it reads: Series A, and in the same way, on your answer sheet, you have a column labeled "A". The problem I am showing you now is problem A1. This means it is the first problem of Series A. Notice what it looks like: the upper part is a figure which is missing a piece [the white space]. Each of the pieces which you find drawn in the lower part of the page [show them] has the correct shape for fitting into this white space, but these pieces cannot complete all of the figure. Number 1 [point to and show the figure to be completed also] provides a completely wrong answer. Numbers 2 and 3 are also not suitable; they fit in the space well, but they do not complete the figure. What do you think, on the other hand, of number 6? It has the correct shape [demonstrate that its shape is identical to that of the preceding numbers] but it does not cover the whole shape. Show me the correct piece (One makes sure that the test participants have found the correct answer. If necessary, supply further explanation, and proceed). *Yes, this is number 4. So the answer to problem A1 is 4. Therefore, write 4 next to the number 1 in column A on your answer sheet. Do not turn the page yet".*

The examiner waits until everyone has finished, and then continues:

"On each page of your booklet, there is a figure which is missing a piece; each time, you have to choose from among the pieces in the lower part, the one which will correctly complete the figure of the upper part. When you have found it, write the corresponding number on your answer sheet, next to the number which indicates the problem's page number. The problems will be easy at the beginning, but they will get steadily more difficult. There are no tricks. If you pay close attention to the way to solve the simple problems, you will find the following ones less difficult.

Solve each problem one after the other, without skipping any of them; start, therefore, at the beginning and continue without pausing until the end. Work at your normal pace. You will not get left behind; you have all the time you need. Now, please turn the page and begin the next problem".

When all the subjects have had time to write their answer for item A2, the examiner will say:

"The correct answer is obviously number 5. Check if you have all written the number 5 next to the number 2 in column A on your answer sheet. Continue on your own now, until the end of the booklet".

At this point, the instructions are complete and the subjects are left free to carry out their work without being pushed in any way, and without providing any further explanations.

– *Test duration*. Subjects are allowed to have all the time they need to complete all the test items. Usually, the average completion time for the SPM is approximately 20 minutes. It is good practice to note the completion time because it can supply important additional information. If the subject terminates in 10 minutes or less, his/her speed can be interpreted as an expression of rapid and concise thought processes, when his/her results are average or higher for his/her age; or it can be seen as an expression of anxiety or of an oppositional attitude to the test, when the results are less than average for his/her age. If the subject finishes after more than 45 minutes, there are two interpretations. If the performance is appropriate to his/her age, the subject may be one whose intelligence is of an analytical nature, which needs processing time to reach the correct answer (from an analysis of the handwriting in the response, should there be marks gone over multiple times or dotted or unsteady ones, one could hypothesize anxious or obsessive traits). If the results are below the average for his/her age and the answers were often erased, corrected, or missing, the

subject may have high performance anxiety and confused or contradictory thinking.

Finally, as far as the choice of when to interrupt testing is concerned, after 45 minutes the subjects are invited to give an answer anyway, and the answer sheets are collected.

– *Data analysis.* SPM's reliability in the Italian standardization sample was measured with the split-half method (correlation between the sums of the even scores and the odd ones) corrected for the length of the test with the Spearman-Brown prophecy formula.

ANOVA were carried out to study the effects of age and gender on the SPM raw scores.

For each of the 13 age range (from 6 to 18 years) of the Italian standardization sample (N = 5438), the SPM raw scores were transformed in standard equivalent points (IQ) with M = 100 and SD = 15, using the continuous norming approach proposed by Gorsuch (Gorsuch, 1983; Zachary & Gorsuch, 1985). This method requires regression analysis procedures which take into account the effect of age on both averages and standard deviations of the raw score distributions. These standards are those referred to by the Giunti Testing report, SPM-RA 6-18 (Standard Progressive Matrices - Report Aggiornato 6-18).

RESULTS

Reliability and standard measurement errors of the SPM

In Table 2, reliability coefficients (r_{tt}) and standard measurement errors (SME) are recorded for each age group, calculated for the SPM raw scores, for each of the ages, with their relative average coefficients. Reliability values vary between .85 (group of 18 year-olds) and .95 (group of 7 year-olds), with an average reliability of .91, which can be considered excellent reliability values. The standard measurement errors tend to increase with age. This means that, contrary to what one might think, the performance of the younger children is more homogenous than that of the oldest young, for which the variability of the "real" score is higher.

Analysis of the effects of age and gender variables on SPM raw scores

First a variation analysis was done with age (13 levels) and gender (2 levels) as independent variables, and raw score obtained on the SPM as the dependent variable. The ANOVA revealed significant differences between the age groups ($F_{12,5412}$ = 539.20; p<.001; η^2 = .54) but not between the two genders ($F_{1,5412}$ = .366; p = .545; η^2 = .00), nor does a significant interaction effect between the two variables emerge ($F_{12,5412}$ = 1.732; p = .054; η^2 = .00); related to the age effect, the post-hoc comparisons (Bonferroni method) between the age groups are generally significant (p<.001), apart from some contiguous age groups (11-12; 12-13; 13-14; 15-16; 17-18).

The trend of the series' average raw scores is the expected one: from the easiest series to the hardest one, results increase with age. Therefore, calculating the percentages of correct responses for age in the various series was possible; they will be taken into consideration in the report.

Analysis of the item difficulty index

The SPM have always been described as being made up of items which are presented in order of difficulty within the series and as being made up of series which get steadily more difficult. However, no one has ever experimentally proven this phenomenon, which has been thought to be true solely

Table 2 – Reliability coefficients and Standard Measurement Errors per age group

Age	6	7	8	9	10	11	12	13	14	15	16	17	18	r_{tt} and average SME
r_{tt}	.91	.95	.91	.93	.92	.92	.86	.91	.88	.89	.90	.90	.85	.91
SME	3.82	3.92	4.60	4.55	4.31	4.16	4.98	4.28	4.95	5.01	5.13	5.22	5.80	4.70

on the basis of the items' content.

The difficulty index has been studied with this sample for Italian standardization, to supply a topic for consideration regarding the single items in each series. In specific, analyzing the percentages of correct answers to the 60 items as a function of age, some irregularities appear. Some items in certain ages are more difficult than items which follow them (e.g. items A8 and A9 at age 6), but those which reveal this irregular trend at all ages are items A11, B8, C4, C8, E8. We could hypothesize, therefore, that when subjects correctly solve one or more of these items (which reveal a low resolving percentage when compared with the items which follow them), this could be interpreted as one of their personal abilities or a strength among their cognitive abilities.

The effect of a social-cultural variable on IQ and the SPM

To study the effect of paternal schooling as an indicator of a family's social-cultural level, on full IQ and the SPM, subjects were divided into 3 levels based on the father's level of schooling (up to 8 years, between 9 and 13 years, and more than 13 years) and an ANOVA was done from which an effect emerges, even if not large, for paternal schooling ($F_{2,5435}$ = 71.218; p<.001; η^2 = .03). Average performance results increase as the level of paternal schooling increases, up to a difference of approximately 7 IQ points between the lower level of schooling and the upper one, while still falling in the average level of interpretation. These values make up one of the standardizing variables in the Giunti Testing SPM report.

Analyzing the trend in the differences among the SPM IQ averages for the subject's age at the time of SPM testing and paternal schooling (see for example Table 3), a linear relationship can be seen between test results increase and paternal schooling increase from 6 years of age, reaching as many as 13 IQ points' difference at age 14. Such differences then tend to decrease after 14 years of age, almost to the point of cancelling themselves out at age 18. Evidently, over time youth's school attendance makes these differences which come from the father's level of schooling less accentuated.

These results confirm what has been found in studies reported in the literature: test performance is correlated to the cultural level of the subject's family (represented in specific by the father's level of instruction) and so it cannot be affirmed that the SPM are a culture-free measure of cognitive capacity. This must be kept in mind when interpreting the results.

Interpretive report. Referring to the SPM-RA 6-18 report by Giunti Testing, after having inserted subjects' responses online, the administrator of the SPM, obtains:

- the percentages of correct answers for each series and for the total, of the test taker and of his/her age group;
- the test taker's total raw score with the corresponding percentile, IQ score, and IQ confidence intervals at 90% and 95%;
- interpretation of the IQ score and confidence interval at 95%;
- interpretation of the test taker's IQ score as a function of the father's years of schooling (if the related personal information field was filled in);
- a general interpretive comment.

Comparison of the English and Italian calibrations

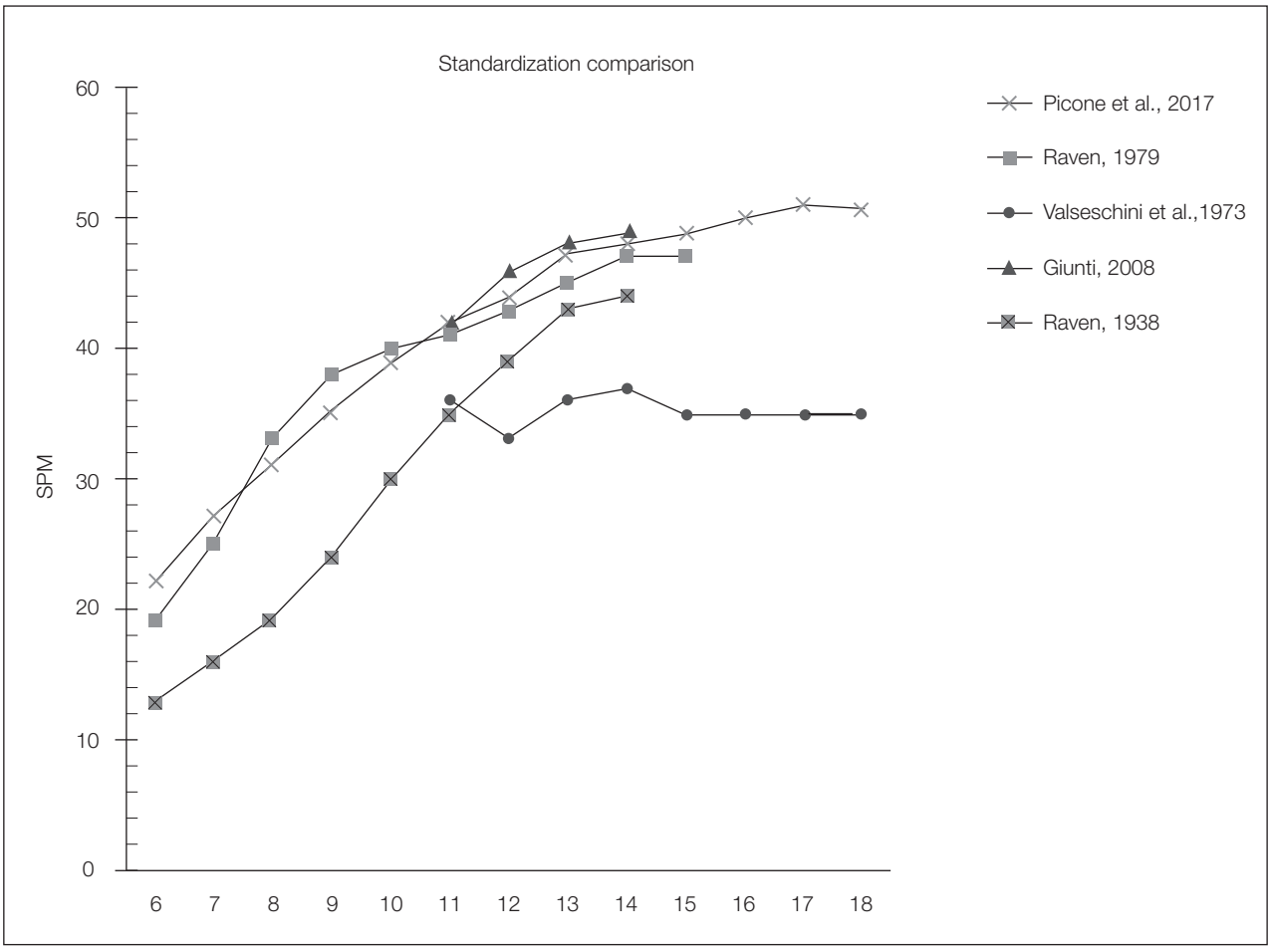
In figure 1 raw score trends at the 50th percentile for each age group are reported, for both the English standardization and the earlier Italian one, to compare them graphically with the current Italian one. The data used for the comparison are taken from the following studies:

- Raven (1938), on 1407 Scottish youth aged 6 to 14;
- Raven (1979), on 3569 Scottish youth aged 6 to 16;
- Valseschini and Del Ton (1973), on 857 subjects aged 11 to 20;

Table 3 – Maximum differences between IQ for subject's age and father's schooling level

Age	6	7	8	9	10	11	12	13	14	15	16	17	18
Maximum difference in IQ points	4	4	6	7	4	6	7	10	13	5	5	5	1

Figure 1 – Comparison of the English and Italian standardizations



– Giunti O.S. (2008), on 825 subjects aged 11 to 14.

As can be seen, the trend our data follows is similar to that of Raven (1979), but, as to that study, our data highlights higher scores starting with age 11. In the most recent Italian study (Giunti O.S., 2008), on a group of young aged 11 to 14, comparable results to ours are found.

The earlier Italian standardization by Valseschini et al, in 1973, beyond offering generally inferior standardization references, included a limited age variation, whereas in the international literature this factor is identified as being the one with the most relevant effect on test performance.

In line with the international literature reports regarding score increase over time for this test according to the so-called "Flynn effect", the differences between the medians found in the first English standardization (Raven, 1938), the subsequent ones (Raven, 1979; Giunti O.S., 2008), and the current Italian standardization are evident in the graph.

CONCLUSIONS

Research which has been published to date underlines how performance obtained on the SPM are correlated with all the cognitive abilities tests, and how the SPM are still used as the best test of abstract, non-verbal reasoning. This ability makes up the essence of the "fluid intelligence" factor, namely the ability to solve logic problems, which stands in contrast to "crystallized intelligence", that being intelligence which uses knowledge people acquire from environmental stimulation and scholastic learning.

Moreover, the kind of fluid intelligence which would be measured displays precisely the problem-solving abilities which are not only present but also potential. It is possible that an individual not be aware of his/her own cognitive abilities, which may not necessarily be expressed in a particularly brilliant scholastic performance. However, if a person is

understood and stimulated, he/she may with time be able to use his/her abilities to progress and to reach important objectives and goals.

The predictive validity regarding scholastic progress at a distance of one or two years is good. Therefore, the SPM can be considered an effective test for predicting the likelihood of scholastic learning, in the absence of serious personality disorders. In this sense, they have been used to predict school success and to select subjects who are well-suited to undertaking university studies, regardless of the disadvantaged environment from which they came.

On the contrary, if individuals do not encounter those people or those things which are able to supply them with adequate motivation, they preserve abilities which remain silent, and bit by bit lose the propulsive thrust which could have lead them to grow and to establish themselves in any context. They are thus flattened in resigned adaptation to prevalent social models and values in their surroundings (above all if these are negative, such as the excessive importance given to consumerism, to establishing a social role, and to scarce interest in knowledge and culture). Thus the importance of the role of attentive, well-prepared educators (be they the parents themselves, teachers, or other adults) who are able to stimulate the energy and abilities present in young people is confirmed.

As to the contrast between hereditary and environmental factors which impact the development of cognitive abilities, the SPM have been considered a culture-free test and, therefore, well-suited to highlighting these abilities regardless of the characteristics of the environment in which the individuals live. The presence of non-verbal stimuli and the limited need for verbal instructions are special features which characterize the SPM as an appropriate test to be administered to subjects who come from others cultures. Currently, the preference is to consider them to be "culture fair" tests, that is "unbiased for all cultures"; they are seen as tests which all individuals can take on, while still being sensible to cultural factors. Research in this area with data gathered from disadvantaged countries compared with more advanced ones which provide greater resources for education, highlights the differences between performance on cognitive ability tests done with urban or rural populations, which have or have not had formal instruction. To close the discussion of the conflict between inheritance and environment, one can consider a review of 30 years of research on ethnic differences in cognitive abilities in which Rushton and Jensen (2005) hypothesize, in agreement with contemporary social sciences, that all children are born with potentially similar intellectual and learning abilities. The inequalities which can be found among various groups are the result of social, economic, and political factors. This world view has generated many strategies for intervention in the family, work, mass media, and criminal justice systems, to the point of including the entire social-economic system. We could think, therefore, that improvement (or worsening) of a population's living conditions can induce a positive (or negative) change in its members' performance on tests for the development of intelligence abilities.

To conclude, the SPM can be used:

- for a quick screening, done collectively, of individuals' level of cognitive efficiency;
- for an assessment of cognitive efficiency in subjects with hearing, language, and motor disabilities;
- for education or professional guidance recommendations (done together with a battery of specific tests);
- regardless of language comprehension problems (e.g., with foreign subjects);
- individually for subjects with comprehension and verbal production deficits or with suspected attention deficits (e.g., subjects diagnosed with ADHD);
- also for subjects with serious personality disorders for whom administration of a tool like the SPM may be useful because it uses abstract and "affectively neutral" stimuli, compared with other cognitive tests;
- for adult and elderly subjects as an indicator of neurophysiological deterioration and as a predictor of degenerative processes (with Alzheimer's, for example).

References

ABDEL-KHALEK, A.M. & RAVEN, J. (2006). Normative data from the standardization of Raven's Standard Progressive Matrices in Kuwait in an international contest. *Social Behavior and Personality*, *34* (2), 169-180.

ADAMS, E.A. (1952). *Analysis of Raven's Matrices scores: Preliminary report*. Surrey, England: Surrey Educational Research Association.

BINET, A. & SIMON, T. (1908). Le développement de l'intelligence chez l'enfant. *L'Année psychologique*, *14*, 1-94.

BOSCHI, F. (1960). Contributo alla taratura del reattivo PM 38 di J.C. Raven. *Bollettino di Psicologia Applicata*, *37*, 173-186.

BURKE, H.R. (1958). Raven's Progressive Matrices: A review and Critical Evaluation. *Journal of Genetic Psychology*, *93*, 199-228.

BYRT, E. & GILL, P.E. (1973). S*tandardisation of Raven's Standard Progressive Matrices and Mill Hill Vocabulary for the Irish population: Ages 6–12*. National University of Ireland, University College Cork.

CHAN, J. (1981). Correlates of parent-child interaction and certain psychological variables among adolescents in Hong Kong. In J.L.M. Binnie-Dawson (Ed.), *Perspectives in Asian cross-cultural psychology* (pp. 117-131). Lisle, Netherlands: Swets and Zeitlinger.

CHAN, J. (1989). The use of Raven's Progressive Matrices in Hong Kong: A critical review. *Psychological Test Bulletin*, *2* (2), 40-45.

de LEMOS, M.M. (1984). A note on the Australian norms for the Standard Progressive Matrices. *Bulletin for Psychologists (ACER Melbourne)*, *36*, 9-12.

de LEMOS, M.M. (1989). The Australian re-standardisation of the Standard Progressive Matrices. *Psychological Test Bulletin, 2* (2), 17-24.

DI FIORE, E. & RENDA, S. (1968). Studio sulle applicabilità di alcuni test in orientamento professionale. *Securitas*, *9-10*, 193-208.

FERRACUTI, F. & GROPPELLI, A. (1954). Contributo preliminare alla taratura del test "Progressive Matrices 1938" di Raven su soggetti di 11, 12, 13 anni di età. *Infanzia Anormale*, *9*, 740-750.

FLYNN, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29-51.

FLYNN, J.R. (1987). Massive IQ gains in 14 nations. What IQ tests really measure. *Psychological Bulletin*, *101*, 171-191.

FLYNN, J.R. (1998). IQ *gains over time: Toward finding the causes*. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measured*. Washington, DC: American Psychological Association.

GIUNTI O.S. (2008). *SPM Standard Progressive Matrices. Standardizzazione italiana*. Firenze: Giunti O.S. Organizzazioni Speciali.

GORSUCH, R.L. (1983). The theory of continuous norming. In R.L. Gorsuch (chair), *Continuous norming: An alternative to tabled norms?* Symposium conducted at the 91ª Annual Convention of the American Psychological Association, Anaheim, CA.

HOLMES, B.J. (1980). *British Columbia norms for Wechsler Intelligence Scale for children-Revised; Peabody Picture Vocabulary Test; Slosson Intelligence Test; Standard Progressive Matrices, and Mill Hill Vocabulary Scale*. Unpublished manuscript, University of British Columbia: Faculty of Education.

JAWOROWSKA, A. & SZUSTROWA, T. (1991). *Podrecznik Do Testu Matryc Ravena*. Warsaw: Pracownia Testow Psychologicznych Polskiego Towarzystwa Psychologicznego.

KAPLAN, R.G. & SACCUZZO, D.P. (1997). *Psychological testing: Principles, applications and issues*. Pacific Grove (CA): Brooks Cole Publishing Company.

KRATZMEIER, H. & HORN, R. (1979). *Manual: Raven-Matrizen-Test, Standard Progressive Matrices*. Weinheim, Germany: Beltz Test.

LYNN, R., ALLIK, J. & IRWING, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, *32*, 411-424.

MACKINTOSH, N.J. & BENNETT, E.S. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. *Intelligence*, *33*, 663-674.

MARTINOLLI, L. (1990). *Étude et Réétalonnage des Matrices Progressives Couleur*. Unpublished manuscript, L'Institut de Psychologie, Université de Fribourg.

OTIS, A.S. (1936-39). *Otis Quick-Scoring Mental Ability Test*. New York: Harcourt, Brace & World.

PICONE, L. (1996). *Il pensiero formale nell'adolescente. Studio metrico della Scala ECDL di Longeot-Hornemann*. Padova: CEDAM.

RAVEN, J.C. (1938). *Progressive Matrices*. London: Lewis.

RAVEN, J.C. (1941). Standardization of Progressive Matrices 1938. *British Journal Medical Psychology*, *19*, 137-150.

RAVEN, J. (1981). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research supplement no. 1: The 1979 British standardization 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*. Oxford, England: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.

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

RAVEN, J. et al. (1990/2000). *Manual for Raven's Progressive Matrices*

and Vocabulary Scales. Research supplement no. 3 (2nd/3rd 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. Oxford, England: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.

RAVEN, J.C., COURT, J.H. & RAVEN, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales. (Section 3) Standard Progressive matrices*. London: Lewis.

RAVEN, J.C., COURT, J.H. & RAVEN, J. (1986). *Manual for Raven's Progressive Matrices and Vocabulary Scales. (Section 3) Research supplement n.3*. London: Lewis.

RAVEN, J.C., COURT, J.H. & RAVEN, J. (1992). *Standard Progressive Matrices, 1992 Edition*. Oxford: Oxford Psychologists Press.

RAVEN, J.C., COURT, J.H. & RAVEN, J. (1995). *Raven, Matrices Progresives (Escalas: CPM, SPM, APM)*. Madrid: TEA Ediciones, S.A.

RAVEN, J. & RAVEN, J. (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. (1978). *Manual for Raven's progressive matrices and vocabulary scales*. London: Lewis

RAVEN, J., RAVEN, J.C. & COURT, J.H. (1998). *Raven manual: Section 3, Standard Progressive Matrices*. Oxford, England: Oxford Psychologists Press.

RAVEN, J.C. & SUMMERS, B. (1986). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research Supplement No. 3*. London: Lewis.

RAVEN, J.C. & WALSHAW, J.B. (1944). Vocabulary tests. *British Journal of Medical Psychology*, *20*, 185-194.

REDA, G.C., NENCINI, R. & RICCIO, D. (1955). Contributo alla taratura delle Progressive Matrices 1937 e 1938 di J.C. Raven. *Bollettino di Psicologia Applicata*, *7-8*, 15-22.

RUSHTON, J.P. & JENSEN, A.R. (2005). Thirty years of research on race differences in cognitive Ability. *Psychology, Public Policy and Law*, *11* (2), 235-294.

SACCUZZO, D.P. & JOHNSON, N.E. (1995). Traditional psychometric test and proportionate representation: An intervention and program evaluation study. *Psychological Assessment*, *7* (2), 183-194.

SPEARMAN, C. (1927). *The abilities of man*. New York: McMillan.

TERMAN, L.M. (1916). *The measurement of Intelligence*. Boston: Houghton Mifflin.

TUDDENHAM, R.D., DAVIS, L., DAVISON, L. & SCHINDLER, R. (1958). *An experimental group version for school children of the Progressive Matrices*. Unpublished manuscript, University of California. (See also Abstract: *Journal of Consultant Psychology*, *22*, 30.)

UNITED STATES GOVERNMENT, BUREAU OF THE CENSUS (1984). S*tatistical Abstract of the United States*, 1983. Washington, DC: U.S. Government Printing Office.

VALSESCHINI, S. & DEL TON, F. (1973). *Le Matrici Progressive di Raven. Contributo critico alla taratura su 1123 soggetti e considerazioni sulla validità e attendibilità della prova*. Firenze: O.S. Organizzazioni Speciali.

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.

WECHSLER, D. (1939). T*he measurement of adult intelligence*. Baltimore: Williams & Wilkins.

ZACHARY, R.A. & GORSUCH, R.L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology*, *41* (1), 86-94.

ZHANG, H.C. & WANG, X.P. (1989). Chinese standardisation of Raven's Standard Progressive Matrices. *Psychological Test Bulletin*, *2* (2), 36-39.