Revista Portuguesa de Investigação Comportamental e Social 2021 Vol. 7(1): 84–101
Portuguese Journal of Behavioral and Social Research 2021 Vol. 7(1): 84–101

ORIGINAL ARTICLE

# Psychometric properties of the Raven's Standard Progressive Matrices in a Portuguese sample

## Propriedades Psicométricas da Forma Geral das Matrizes Progressivas de Raven numa amostra portuguesa

**Inês Queiroz-Garcia** (1)
**Helena Espirito-Santo** (2,3)
**Catarina Pires** (4)

(1) *Inst Universitário (ISPA), William James Center for Research, Portugal*
(2) *Inst Superior Miguel Torga, Coimbra, Portugal*
(3) *Univ Coimbra, Fac of Psychology and Educational Sciences; Centre for Research in Neuropsychology and Cognitive and Behavioural Intervention (CINEICC), Portugal*
(4) *Independente, Aveiro, Portugal*

Recebido: 15/04/2021; Revisto: 17/05/2021; Aceite: 18/05/2021.

https://doi.org/10.31211/rpics.2021.7.1.210

Check for updates

## Abstract

**Objective:** Psychometric properties of Raven's Standard Progressive Matrices in a Portuguese community sample were investigated. **Method**: The sample consists of 522 people (250 men and 272 women), aged between 12 and 95 years. All participants completed an informed consent form and a battery of neuropsychological tests, including Raven's Standard Progressive Matrices (RSPM), Rey 15-Item Memory Test, Zung Self-Rating Anxiety Scale, and Rey Complex Figure Test. **Results**: The average in RSPM was 41.18 (*SD* = 12.03). The results showed that all sociodemographic variables (age, sex, education, profession, regions, and place of residence) significantly influenced RSPM scores. The reliability and temporal stability of RSPM were adequate. **Conclusions**: This study suggests that RSPM is an instrument with potential for use among the Portuguese population.

**Keywords:** Intelligence; Neuropsychological tests; Portuguese population; Raven's Standard Progressive Matrices; Psychometric study.

**DI&D | ISMT**

rpics@ismt.pt

https://rpics.ismt.pt



Helena Espírito Santo

Rua Augusta, 46
3000-061
Coimbra, Portugal,
E-mail: espirito-santo@ismt.pt

## Resumo

**Objetivo**: Foram investigadas as propriedades psicométricas da Forma Geral das Matrizes Progressivas de Raven numa amostra da comunidade portuguesa. **Métodos**: Participaram 522 pessoas (250 homens e 272 mulheres), com idades compreendidas entre os 12 e os 95 anos. Todos os participantes preencheram um formulário de consentimento informado e uma bateria de testes neuropsicológicos, incluindo o teste da Forma Geral das Matrizes Progressivas de Raven (FGMPR), Teste de Memória de 15-Item de Rey, Escala de Ansiedade Zung Self-Rating e Teste de Figura Complexa de Rey. **Resultados**: A média na FGMPR foi de 41,18 (*DP* = 12,03). Os resultados mostraram que todas as variáveis sociodemográficas (idade, sexo, educação, profissão, regiões e local de residência) influenciaram significativamente os resultados da FGMPR. A fiabilidade e estabilidade temporal da FGMPR foram adequadas. **Conclusão**: Este estudo sugere que a FGMPR é um instrumento com potencial para ser utilizado na população portuguesa.


**Palavras-Chave:** Inteligência; Testes neuropsicológicos; População portuguesa; Forma Geral das Matrizes Progressivas de Raven; Estudo psicométrico.

## Introduction

The concept of *g* or general intelligence was created by Charles Spearman, who established *g* as a measure of performance in a variety of tests (Carter, 2005). Spearman's theory is based on two factors: the *g* factor (general intelligence) and the *s* factor (specific intellectual abilities). Spearman, however, implies that a real intelligence test will be highly *g* loaded (Murphy & Davidshofer, 2004; Simões, 2000). Several studies contend that Raven's Standard Progressive Matrices (RSPM) is one of the purest and best measures of the *g* factor (Raven et al., 2000; Jensen, 1980; 1998; Orme, 1966). The *g* factor is based on two components: eductive ability and reproductive ability. The first corresponds to the capacity to forge new knowledge, discern meaning in confusion, perceive, and identify relationships. The RSPM was conceived to measure this eductive ability (Raven et al., 2000). Additionally, Cattell (1963) suggests two related but conceptually distinct aspects: fluid intelligence and crystallized intelligence. Fluid intelligence is defined as the ability to see relationships, dealing mainly with reasoning capability (Cairo Martínez et al., 2000; Cattell, 1963; Murphy & Davidshofer, 2004). Fluid intelligence is most loaded in tests with virtually no educational or cultural content, such as RSPM (Jensen, 1998). Therefore, this test allows assessing abstract thinking capabilities, solving new problems, reasoning, and analogies (Cairo Martínez et al., 2000). On another classification, Jensen (1987) raised the hypothesis of two types of entirely different cognitive processes, which he called Level I and Level II. Level II's ability involves transformation or manipulation of the input to arrive at the output through abilities like reasoning, problem-solving, inference, generalization, and conceptual categorization. Level II is much the same as Spearman's *g*, common to all complex cognitive abilities' tests. Similarly, Level II is near what Cattel referred to as fluid intelligence (Jensen, 1987). Jensen (1998), Kaplan and Saccuzzo (2012) consider that RSPM is a fluid reasoning measure. In all the theories, reasoning seems to be the common element (Simões, 2000).

The best-known form of the Raven's Progressive Matrices (RPM) (published in 1938) is RSPM, devised in England by J. C. Raven, a psychologist, and L. S. Penrose, a geneticist. It was created to measure as completely as possible in a single test the eduction processes that Spearman regarded as the essence of intelligence (Jensen, 1980). The RSPM is the most widely used test on five continents (Irvine & Berry, 1988), being suitable from age six to adulthood since it covers a greater cognitive ability variation (Jensen, 1980; Raven, 1941). Orme (1966), Jensen (1980), Karnes et al. (1982) consider that the RSPM measure a very fundamental aspect of cognitive performance not influenced by cultural influences. RSPM consists almost entirely of perceiving key features and relationships and discovering the abstract rules that govern the differences among the elements in the matrix. The RSPM test consists of a set of entirely non-verbal tasks specially designed to reduce item dependence on acquired knowledge, cultural and scholastic content while calling for reasoning ability (Jensen, 1980).

The RSPM consists of sixty diagrammatic puzzles divided into five sets (A, B, C, D, and E) (Lovett et al., 2007; Raven et al., 2000). The items are organized in/sets of gradual difficulty: each set begins with easy items, and these are followed by items that become progressively more challenging (Jensen, 1980; Raven, 1941; Raven et al., 2000). Each item has one missing cell in a matrix (always in the lower right corner), and the subject is required to select the best completion solution among a number of choices (Jensen, 1980; Lovett et al., 2007; Raven, 1941; Raven et al., 2000). The solution depends on the complexity of the items (abstract material through figures or geometric designs) and the number of elements involved in reasoning ability. The variety of forms, relationships, and transformations are practically unlimited (Jensen, 1980). The five sets provide the opportunity to understand the method of thought required to solve the problems and five progressive assessments of a person's intellectual ability capacity (Raven et al., 2000). To ensure interest and the absence of fatigue, each problem is presented and elaborated with precision and as far as possible, pleasing to look at, ensuring that the person does not become discouraged and lose interest in the subject. The test instructions are so simple, and the tasks requirements so obvious to most of the participants that the instructions may even be given by pantomime (Jensen, 1980; Raven et al., 2000).

The RSPM can be individually or to a group (Jensen, 1980). The RSPM manual stipulates that, if time is limited, then the test should be used as a measure of *intellectual efficiency* and not the *g* factor (Raven et al., 2000). The timed version is not a pure measure of fluid ability given the fact that it is influenced by a speed factor that discriminates against people who work more slowly and carefully (Raven et al., 2000).

Despite the popularity of RSPM, one limitation is the time required to administer the full set of items (Bilker et al., 2012; Chiesi et al., 2012). The main concern regarding its length is the possibility of excluding some items, which might limit the differentiation of different levels of performance. The failure to distinguish between the various levels of performance and the diminution of administration time results in an incapability to maintain the original scale's properties, especially its discriminative power (Chiesi et al., 2012). For example, a correct answer for some of the RSPM items requires the simultaneous mental manipulation of three or four different fundaments, which takes some time. The easier and much less time-consuming manipulation of only one or two of these fundaments will lead to a reliable selection of one of the distractors, which is generally seen in cognitively less able participants. Participants who have enough time to solve each matrix come to conclusions that it is not based on pure guesswork (even when wrong). Furthermore, the additional time is of little use (Jensen, 1980).

The RSPM can be applied in school contexts, workplaces, and even at home, as well as in the laboratory (Raven et al., 2000). Due to its non-verbal format, the RSPM is used with a broad range of populations for whom language processing is minimized, for example, foreign language groups, the deaf, and also with economically disadvantaged groups (Jensen, 1980; Karnes et al., 1982; Raven et al., 2000). The RSPM has also been used in several cross-cultural studies (Jensen, 1980; Karnes et al., 1982; Kush, 1996; Powers & Barkan, 1986).

In fact, the literature review shows that the RSPM is a viable and promising tool in several cultures, including the Egyptian (Abdel-Khalek, 1988), Irish (Moran, 1986), Libyan (Al-Shahomee, 2012), Kuwait Chileans older adults (Alarcón Paz et al., 2012), in an African sample (Rushton et al., 2004), Hispanic Anglo-American (Powers

& Barkan, 1986), and in Icelandic children (Pind et al., 2003). It has also been useful in studies with computerized versions equivalent to RSPM (Golan et al., 2019; Williams & McCord, 2006) and abbreviated versions (Bilker et al., 2012; Irrgang et al., 2019).

We chose the RSPM because it is the most widely used test compared to the Special Form and Advanced Form of the MPR (Raven & Raven, 2003).

This study's primary goal was to identify the psychometric properties of the RSPM in a sample of the Portuguese population since the psychometric data of RSPM is lacking. As specific goals, we intended to explore (1) the effect of variables such as age, sex, and education in the performance of RSPM, (2) to determine the internal consistency, and the test-retest reliability, (3) the convergent validity, correlating RSPM to other instruments used in the test battery.

## Methods

### Participants

A sample of 525 volunteers was recruited using a convenience sampling method. Recruitment took place in the community, and 1.1% (*n* = 6) refused to participate. People did not receive any financial compensation for participating but were given the option to receive the results and their interpretation (64 participants requested these results, 12.3%).

Predetermined criteria that precluded consideration for the study included (a) being of Portuguese nationality or living in Portugal for more than five years, (b) having received more than 50% of their education in Portugal.

The participants were stratified according to age: 12 to 17 years; 18 to 30 years; 31 to 40 years; 41 to 50 years; 51 to 60 years; 61 to 70 years; 71 to 80 years and 81 to 90 years.

Participants in our sample were also stratified according to their educational level, divided into four categories: without education, basic education (we have grouped the three basic education levels), secondary education, and higher education[1]. We have no information on thirty-five of the participants.

Portuguese regions were grouped according to our territorial units (North, Central Region, South, and Autonomous Regions). In the South region, we included Alentejo and Algarve and merged with Autonomous Regions for statistical analysis. We have no information on 162 of the participants.

The residence place was evaluated by questioning whether the participants lived in an urban, transition area, or a rural environment (Statistics Portugal, 2014). We have no information on forty-two of the participants.

### Procedure and Measures

All participants signed a consent form which included a description of the experiment, rights to confidentiality, and their right to end participation according to the Helsinki Declaration. All the participants were evaluated individually by five graduate students trained in administering the tests with an approximate duration of one hour. Most of the participants were administered the Clock Drawing Test (Shulman et al., 1986), Rey 15-Item Memory

[1] In the Portuguese education system the 1st cycle of basic education is from 1st to 4th grade; the 2nd cycle of basic education is from 5th to 6th grade; 3rd basic education is from 7th to 9th grade; secondary education is from 10th to 12th grade, and the higher education corresponds to the university or college.

Test (Boone et al., 2002), Zung Self-Rating Anxiety Scale (Serra et al., 1982), Edinburgh Handedness Inventory (Espírito-Santo et al., 2017), Frontal Assessment Battery (Dubois et al., 2000), Stroop test (Garcia et al., 2016; Stroop, 1935), RSPM (Raven et al., 2000), and the Rey Complex Figure Test (Rey, 2002). In the RSPM, all subjects were instructed by the guidelines established by the authors of the test (Raven et al., 2000).

RSPM (Raven et al., 2000), a non-verbal test, consists of 60 tasks presented in black and white and divided into five series numbered alphabetically. Each set consists of twelve items. Participants were asked to select the missing part from six or eight options given below each matrix. A participant's score is the number of correct answers, ranging from 0 to 60 (Al-Shahomee, 2012; Bilker et al., 2012; Raven & Raven, 2003; Savage-McGlynn, 2012). Each participant completed the RSPM hetero-administrative form. The RSPM was administered untimed (Raven et al., 2000). Administration time varied from 9 min to 120 min, and the average administration time was 32.32 (*SD* = 15.11).

#### *Rey 15-Item Memory Test (15-IMT)*

The 15-IMT (Boone et al., 2002) evaluates suspect effort of participants. The 15-IMT involves the memorization of 15 different items, making it appear the task harder than it is. The test consists of two tasks: recall task and recognition task. This formula provides the score: recall correct + (recognition correct – false positive). Given the result of the application of this formula, consider a cut of 20 points. We used the test to detect the simulation and the insufficient effort of the participants (Griffin et al., 1997; Reznek, 2005; Simões et al., 2010; Slick et al., 2004). The Cronbach's alpha of this instrument was .80, and 174 people responded to this tool because of the moment the 15-IMT was included in the neuropsychological battery.

#### *Zung Self-Rating Anxiety Scale (SAS Zung)*

The SAS Zung (Serra et al., 1982) is a self-administered scale to measure anxiety state; namely, the anxiety felt at the time of the evaluation. This scale consists of twenty items on a Likert scale where items responses are ranked from 1 to 4, except the items 5, 9, 13, 17, 19, which are scored in reverse (4, 3, 2, 1). The total score ranges from 20 to 80, with higher scores corresponding to more frequent symptoms. Results above 40 indicate that the individual has a high likelihood of suffering from clinically significant anxiety. We used this test to discriminate between participants who experienced anxiety at the time of the evaluation of those who did not (Serra et al., 1982). The Cronbach's alpha of this instrument was .78. One hundred seventy-four people answered to SAS Zung, also because of its timing on the neuropsychological battery.

#### *Rey Complex Figure Test (RCFT)*

The RCFT (Rey, 2002) consists of three tasks that occur at different times: the copy task, the immediate recall (3 minutes), and delayed recall (20 minutes). In the copy task, the participant must reproduce the figure on a piece of paper. After three minutes, the participant is asked to reproduce the figure from memory without prior warning. Twenty minutes later, the participant is requested to reproduce the figure from memory again. Each copy is scored for the accurate reproduction and placement of 18 elements and may have a maximum score of 36 points (Rey, 2002). This test evaluates implicit memory and visuospatial ability, and we have selected this test because

 DOI: 10.31211/rpics.2020.7.1.210

Thurstone (Murphy & Davidshofer, 2004) had considered these skills as primary mental abilities. The Cronbach's alpha of this instrument was .85 ($n$ = 199).

### Statistical Analysis

We used the Statistical Package for the Social Sciences (IBM SPSS Statistics for Macintosh, Version 27.0) to perform data analysis.

The analysis was carried out in the following manner. First, a descriptive statistical analysis was used to determine the mean, standard deviation, frequencies, percentages of RSPM scores. To explore the proportion of cases that fall into each category of each demographic variable categorized and compare this ratio with hypothetical values, we used a chi-square test for goodness of fit (Pallant, 2016).

Then it was verified that the normality of score distribution of RSPM using the Shapiro-Wilk test and the measures of kurtosis and skewness.

We used Student's $t$-test/ANOVA of one and two factors and the Pearson correlations to explore the effects of sociodemographic characteristics (age, sex, education, regions, and place of residence) on the performance of RSPM. For the independent samples $t$-test, a power of .95, and an alpha of .05, the sample would need 210 participants. As for the ANOVA of one factor, a .95 power, and a .05 alpha, the sample would have between 252 and 360 attendees. As for the two-factor ANOVA, .95 and power a .05 alpha, the sample would need to have 400 participants (Faul et al., 2007a, 2007b). For the analysis of variance (ANOVA), we determined the homogeneity of variances according to Levene's test. In the case of existing homogeneity ($p > .05$), we used the Hochberg posthoc test. Otherwise, the Games-Howell posthoc test would be used both with the Bonferroni correction ($p$ / number of pairwise comparisons).

For the analysis of the psychometric properties, we determined the internal consistency through Cronbach's alpha. For the analysis of temporal stability test-retest, we used Pearson's correlations and paired samples $t$-test. For the analysis of internal consistency for an alpha of .01, a power of .95, and an expected Cronbach's alpha of .89 (Moran, 1986), the sample size needed would be 13 participants (Bonett, 2002; Chang, 2014). The test's reliabilities for the different groups were determined by recoding the items to 0 (incorrect) and 1 (correct).

For convergent validity, we computed correlational analysis with some of the tests used in the neuropsychological battery. As for the correlational analysis, a power of .95 and an alpha of .05, the sample would have to consist of 314 participants (Faul et al., 2007a, 2007b). We also computed correlational analysis relating the RSPM scores with age, years of study, and total time took by the participants.

## Results

From the initial pool (525 participants), three people (two Brazilian and one Angolan nationals) were excluded since our study's goal was to confine it to the Portuguese population.

Seventeen participants that scored below 19 on 15-IMT test had significantly lower scores on RSPM [participants with suspect effort: $M \pm SD$ = 28.35 ± 11.41; non-suspect effort: $M \pm SD$ = 43.00 ± 9.56; $t_{(172)}$ = 5.89;

 DOI: 10.31211/rpics.2020.7.1.210

$p < .001$; $CI$95% [8.98; 20.32], than those of the participants without suspect effort. The effect size was very large (Cohen's $d = 1.50$; $CI$95% [0.97; 2.02]).

Forty participants with Zung scores below 40 performed poorly on RSPM [anxious: $M \pm SD = 38.00 \pm 12.21$; non-anxious participants: $M \pm SD = 42.63 \pm 9.95$; $t_{(55.358)} = 2.19$; $p < .05$; $CI$95% [0.46; 8.80], but the effect size was small (Cohen's $d = 0.44$; $CI$95% [0.08; 0.79]).

Of these 522 participants (Table 1), 250 were men (47.9%) and 272 women (52.1%). Ages ranged between 12 and 95 years ($M = 46.35$; $SD = 20.92$).

The educational level ranged between illiteracy (3.1%) and 28 years of schooling ($M = 10.23$; $SD = 5.30$). According to the territorial units, 333 subjects (81.4%) lived in the Central region (Aveiro, Castelo Branco, Coimbra, Guarda, Leiria, part of Santarém, and Viseu) 41 (10.0%) in the North, 16 (3.9%) in Lisbon Metropolitan Area (Lisboa, Sintra), 15 (3.7%) in Alentejo and four (1.0%) in the autonomous regions of Azores and Madeira.

Regarding the place of residence, 257 subjects (51.7%) lived in an urban area, while 194 (39.0%) and 46 (9.3%) lived in a rural area and a transition area, respectively.

Through the chi-square test for goodness of fit is possible to verify that the sample was not balanced by sociodemographic variables.

**Table 1**

*Sociodemographic Characteristics*

| | | *n* | % | Census 2011 |
|---|---|---|---|---|
| Age | | | | |
| ($M \pm SD = 46.35 \pm 20.92$) | 10-19 | 59 | 11.3 | 10.7 |
| | 20-29 | 70 | 13.4 | 11.7 |
| | 30-39 | 85 | 16.3 | 15.1 |
| | 40-49 | 83 | 15.9 | 14.6 |
| | 50-59 | 80 | 15.3 | 13.3 |
| | 60-69 | 67 | 12.8 | 11.2 |
| | ≥ 70 | 78 | 14.9 | 13.8 |
| Sex | | | | |
| | Men | 250 | 47.9 | 47.8 |
| | Women | 272 | 52.1 | 52.2 |
| Education level | | | | |
| ($M \pm SD = 10.23 \pm 5.30$) | Without education | 15 | 3.1 | 8.3 |
| | Basic education | 256 | 52.6 | 54.7 |
| | Secondary education | 109 | 22.4 | 19.9 |
| | Higher education | 107 | 22.0 | 17.1 |
| Regions | | | | |
| | North | 41 | 10.0 | 35.1 |
| | Center | 349 | 85.3 | 21.9 |
| | South and A.R. | 19 | 4.6 | 16.3 |
| Area of residence | | | | |
| | Urban | 257 | 51.7 | — |
| | Transition area | 46 | 9.3 | — |
| | Rural | 194 | 39.0 | — |

*Note*. $N = 522$. South and A.R. = South and Autonomous Regions (Azores and Madeira).

 DOI: 10.31211/rpics.2020.7.1.210

## Descriptives

The average RSPM score was 41.18 ($SD$ = 12.03). Shapiro-Wilk test indicated a non-normal distribution ($p$ < .001). The RSPM scores distribution was negative and slightly skewed ($Si$ = -0.78) and slightly leptokurtic ($Ku$ = -0.13). However, kurtosis values (lower than 7, proper value) and the skewness (over -2), according to Kim (2013), is indicative of normality for samples exceeding 300 participants, allowing the use of parametric tests to investigate and evaluate the presence of statistically significant differences in data.

## Influence of Sociodemographic Variables

Table 2 presents the differences in the scores of RSPM between groups defined by sociodemographic variables, using Student's *t*-test for independent samples or analysis of variance (ANOVA).

The total scores on RSPM differed significantly according to seven age groups [$F_{(6, 468)}$ = 41.91; $p$ < .001], with a large effect ($h^2$ = 0.33) (Table 2). The correlational analysis confirms the trend of a decrease in RSPM score relative to age, showing, according to the criteria of Cohen (1988), a high negative correlation ($r$ = -.53; $p$ < .001).

Regarding sex (Table 2), men had statistically significant higher scores than women [$t_{(516.080)}$ = 3.92; $p$ < .001]. Nevertheless, the effect size was small (Cohen's $d$ = 0.34).

The total scores on RSPM differed significantly according to educational levels [$F_{(3, 195)}$ = 85.75; $p$ < .001], with a large effect ($h^2$ = 0.27) (Table 2). The correlational analysis confirmed the trend of an increased score of RSPM through years of study, showing, according to the criteria of Cohen (1988), a high positive correlation ($r$ = .57; $p$ < .001).

The total scores on RSPM (Table 2) differed significantly between regions [$F_{(2, 74)}$ = 10.12; $p$ < .001], with a small effect ($h^2$ = 0.03).

The total scores on RSPM differed significantly between the area of residence [$F_{(2, 140)}$ = 9.37; $p$ < .001], with a small effect ($h^2$ = 0.04) (Table 2).

Games-Howell posthoc tests (Table 3), with Bonferroni correction, revealed that the five age groups between 10 and 49 years had significantly higher scores than those groups above 60 years. Also, those in the 50 – 69 age groups differed from those aged above 70 years.

Games-Howell posthoc test (Table 4), with Bonferroni correction, revealed that participants' scores without education differed significantly from participants with all other education levels. Similarly, the scores of participants with secondary and higher education differed significantly from participants who had basic education.

Games-Howell posthoc test (Table 5), with Bonferroni correction, revealed that scores of participants who resided in the Center region were significantly higher than those in the North.

The Games-Howell posthoc test (Table 6) with Bonferroni correction revealed that scores of participants who resided in an urban or rural area were significantly higher than those residing in a transition zone.

 DOI: 10.31211/rpics.2020.7.1.210

**Table 2**

*Differences in Scores of Raven's Standard Progressive Matrices in the Sociodemographic Variables*

| | | *n* | *M* ± *SD* | *CI*95% *IL* – *UL* | Min – Max |
|---|---|---|---|---|---|
| Age | | | | | |
| $F_{(6, 468)}$ = 41.91; | 10 – 19 | 59 | 45.83 ± 7.28 | 43.93 – 47.73 | 24 – 59 |
| $p$ < .001; $\eta^2$ = 0.33 | 20 – 29 | 70 | 46.90 ± 7.84 | 45.03 – 48.77 | 25 – 60 |
| | 30 – 39 | 85 | 46.20 ± 9.32 | 44.19 – 48.21 | 18 – 60 |
| | 40 – 49 | 83 | 44.39 ± 9.92 | 42.22 – 46.55 | 18 – 60 |
| | 50 – 59 | 80 | 42.48 ± 10.84 | 40.06 – 44.89 | 12 – 59 |
| | 60 – 69 | 67 | 35.87 ± 11.97 | 32.95 – 38.78 | 9 – 58 |
| | > 70 | 78 | 26.90 ± 11.09 | 24.40 – 29.40 | 2 – 51 |
| Sex | | | | | |
| $t_{(516.080)}$ = 3.92; | Men | 250 | 43.29 ± 10.77 | 41.95 – 44.63 | 9 – 60 |
| $p$ < .001; $d$ = 0.34 | Women | 272 | 39.25 ± 12.80 | 37.72 – 40.77 | 2 – 60 |
| Education level | | | | | |
| | Without education | 15 | 20.93 ± 7.21 | 16.94 – 24.92 | 14 – 36 |
| $F_{(3, 195)}$ = 85.75; | Basic education | 256 | 36.96 ± 12.01 | 35.48 – 38.43 | 2 – 59 |
| $p$ < .001; $\eta^2$ = 0.27 | Secondary education | 109 | 46.69 ± 8.05 | 45.16 – 48.22 | 24 – 60 |
| | Higher education | 107 | 47.91 ± 8.07 | 46.36 – 49.45 | 12 – 60 |
| Regions | | | | | |
| $F_{(2, 74)}$ = 10.12; | North | 41 | 34.83 ± 11.24 | 31.28 – 38.38 | 15 – 59 |
| $p$ < .001; $\eta^2$ = 0.03 | Center | 349 | 41.79 ± 12.11 | 40.52 – 43.07 | 12 – 60 |
| | South and A.R. | 19 | 41.42 ± 6.55 | 38.26 – 44.58 | 30 – 55 |
| Area of residence | | | | | |
| $F_{(2, 140)}$ = 9.37; | Urban | 257 | 42.65 ± 11.36 | 41.25 – 44.04 | 12 – 60 |
| $p$ < .001; $\eta^2$ = 0.04 | Transition area | 46 | 33.93 ± 13.83 | 29.83 – 38.04 | 13 – 60 |
| | Rural | 194 | 41.03 ± 11.56 | 39.39 – 42.67 | 9 – 59 |

*Note. N* = 522. *d* = Cohen's d; $\eta^2$ = eta-squared (sum of squares between groups / sum of the squares); South and A.R. = South and Autonomous Regions.

**Table 3**

*Posthoc Comparisons of Scores of Raven's Standard Progressive Matrices in Age Groups*

| Pairwise comparisons[a] | | Mean difference | *p* | *d* | Interpretation *d* |
|---|---|---|---|---|---|
| 10 – 19 | 60 – 69 | 9.97 | < .001 | 0.99 | large effect |
| (*M* ± *SD* = 45.83 ± 7.28) | > 70 | 18.93 | < .001 | 1.58 | very large effect |
| 20 – 29 | 60 – 69 | 11.03 | < .001 | 0.92 | large effect |
| (*M* ± *SD* = 46.90 ± 7.84) | > 70 | 20.00 | < .001 | 1.67 | very large effect |
| 30 – 39 | 60 – 69 | 10.33 | < .001 | 0.86 | large effect |
| (*M* ± *SD* = 46.20 ± 9.32) | > 70 | 19.30 | < .001 | 1.61 | very large effect |
| 40 – 49 | 60 – 69 | 8.52 | < .001 | 0.71 | medium effect |
| (*M* ± *SD* = 44.39 ± 9.92) | > 70 | 17.49 | < .001 | 1.46 | very large effect |
| 50 – 59 (*M* ± *SD* = 42.48 ± 10.84) | > 70 | 15.58 | < .001 | 1.30 | very large effect |
| 60 – 69 (*M* ± *SD* = 35.87 ± 11.97) | > 70 | 8.97 | < .001 | 0.75 | medium effect |

*Note*. *N* = 522. $F_{(6, 468)}$ = 41.91; *p* < .001. *d* = Cohen's *d*.

[a] Only statistically significant differences are presented.


DOI: 10.31211/rpics.2020.7.1.210

**Table 4**

*Posthoc Comparisons of Scores of Raven's Standard Progressive Matrices in Educational Levels*

| Pairwise comparisons[a] | | Mean difference | *p* | *d* | Interpretation *d* |
|---|---|---|---|---|---|
| Basic education | Without education | 16.02 | < .001 | 1.34 | very large effect |
| Secondary education | Without education | 25.76 | < .001 | 2.15 | very large effect |
| ($M \pm SD$ = 46.69 ± 8.05) | Basic education | 9.73 | < .001 | 0.81 | large effect |
| Higher education | Without education | 26.97 | < .001 | 2.25 | very large effect |
| ($M \pm SD$ = 47.91 ± 8.07) | Basic education | 10.95 | < .001 | 0.91 | large effect |

*Note*. $N$ = 522. $F_{(3, 195)}$ = 85.75; $p < .001$. $d$ = Cohen's $d$.
[a] Only statistically significant differences are presented.

**Table 5**

*Posthoc Comparisons of Scores of the Raven's Standard Progressive Matrices in Regions of Residence*

| Pairwise comparisons[a] | | Mean difference | *p* | *d* | Interpretation *d* |
|---|---|---|---|---|---|
| Center | North | 6.96 | < .01 | 0.58 | medium effect |

*Note*. $N$ = 522. $F_{(2, 74)}$ = 10.12; $p < .001$. $d$ = Cohen's $d$.
[a] Only statistically significant differences are presented.

**Table 6**

*Posthoc Comparisons of Scores of the Raven's Standard Progressive Matrices in Place of Residence*

| Pairwise comparisons[a] | | Mean difference | *p* | *d* | Interpretation *d* |
|---|---|---|---|---|---|
| Urban ($M \pm SD$ = 42.65 ± 11.36) | Transition area | 8.71 | < .001 | 0.73 | medium effect |
| Rural ($M \pm SD$ = 41.03 ± 11.56) | Transition area | 7.10 | < .01 | 0.59 | medium effect |

*Note*. $N$ = 522. $F_{(2, 140)}$ = 9.37; $p < .001$. $d$ = Cohen's $d$.
[a] Only statistically significant differences are presented.

A two-way analysis of variance (age groups x sex) showed a non-significant effect on RSPM scores [$F_{(6, 508)} = 1.63$; $p > .05$]. The scores by age groups and sex are presented in Table 7.

A two-way analysis of variance (education level x age groups) showed a significant effect on RSPM scores [$F_{(9, 468)} = 2.45$; $p < .05$]. The scores by education level and age groups are presented in Table 8.

 DOI: 10.31211/rpics.2020.7.1.210

**Table 7**

*Descriptives of Scores of the Raven's Standard Progressive Matrices by Age Groups x Sex*

| Age groups | Men | | Women | |
|---|---|---|---|---|
| | *n* | *M* ± *SD* | *n* | *M* ± *SD* |
| 10 – 19 | 17 | 43.29 ± 8.75 | 42 | 46.86 ± 6.44 |
| 20 – 29 | 53 | 46.64 ± 8.39 | 17 | 47.71 ± 5.92 |
| 30 – 39 | 52 | 47.83 ± 8.98 | 33 | 43.64 ± 9.39 |
| 40 – 49 | 39 | 44.15 ± 10.44 | 44 | 44.59 ± 9.56 |
| 50 – 59 | 39 | 43.97 ± 9.26 | 41 | 41.05 ± 12,10 |
| 60 – 69 | 28 | 36.04 ± 11.73 | 39 | 35.74 ± 12.29 |
| > 70 | 22 | 30.95 ± 10.31 | 56 | 25.30 ± 11.06 |

*Note*. *N* = 522.

**Table 8**

*Descriptives of the Raven's Standard Progressive Matrices Scores by Education Level x Age*

| Education Level | | Age Groups | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 10 – 19 | 20 – 29 | 30 – 39 | 40 – 49 | 50 – 59 | 60 – 69 | > 70 |
| Without education | | | | | | | | |
| | *n* | — | — | — | — | — | — | 15 |
| | *M* | — | — | — | — | — | — | 20.93 |
| | *(SD)* | (—) | (—) | (—) | (—) | (—) | (—) | (7.21) |
| Basic education | | | | | | | | |
| | *n* | 23 | 7 | 18 | 37 | 60 | 52 | 59 |
| | *M* | 47.39 | 39.57 | 42.28 | 39.38 | 40.15 | 33.40 | 29.32 |
| | *(SD)* | (7.17) | (11.57) | (9.46) | (11.18) | (11.29) | (11.06) | (11.03) |
| Secondary education | | | | | | | | |
| | *n* | 26 | 16 | 33 | 15 | 10 | 9 | — |
| | *M* | 44.08 | 47.56 | 46.58 | 47.93 | 49.10 | 48.33 | — |
| | *(SD)* | (7.91) | (7.92) | (9.22) | (8.31) | (4.63) | (6.34) | (—) |
| Higher education | | | | | | | | |
| | *n* | — | 26 | 34 | 31 | 10 | 6 | — |
| | *M* | — | 48.46 | 47.91 | 48.65 | 49.80 | 38.50 | — |
| | *(SD)* | (—) | (7.25) | (8.99) | (5.56) | (5.53) | (15.06) | (—) |

*Note*. *N* = 522.


DOI: 10.31211/rpics.2020.7.1.210

A two-way analysis of variance (education level x sex) revealed a non-significant effect on RSPM scores [$F_{(3, 479)}$ = 1.09; $p$ > .05]. Table 9 presents the scores by education level and sex.

**Table 9**

*Descriptives of the Raven's Standard Progressive Matrices Scores by Education Level x Sex*

| Education Level | Men | | Women | |
|---|---|---|---|---|
| | *n* | *M* ± *SD* | *n* | *M* ± *SD* |
| Without education | 3 | 19.33 ± 9.24 | 12 | 21.33 ± 7.05 |
| Basic education | 114 | 39.32 ± 10.46 | 142 | 35.06 ± 12.84 |
| Secondary education | 60 | 47.65 ± 8.16 | 49 | 45.51 ± 7.85 |
| Higher education | 65 | 48.08 ± 8.07 | 42 | 47.64 ± 8.16 |

*Note*. *N* = 522.

A two-way analysis of variance (age groups x regions) revealed a non-significant effect on RSPM scores [$F_{(10, 390)}$ = 1.78; $p$ > .05] (Table 10).

**Table 10**

*Descriptives of the Raven's Standard Progressive Matrices Scores by Age Groups x Regions*

| Age Groups | North | | Center | | South and A.R. | |
|---|---|---|---|---|---|---|
| | *n* | *M* ± *SD* | *n* | *M* ± *SD* | *n* | *M* ± *SD* |
| 10 – 19 | 8 | 41.00 ± 6.33 | 41 | 47.46 ± 7.19 | 6 | 40.33 ± 5.16 |
| 20 – 29 | 7 | 41.29 ± 10.11 | 51 | 47.65 ± 7.89 | 2 | 48.00 ± 9.90 |
| 30 – 39 | 5 | 41.20 ± 14.64 | 51 | 46.35 ± 9.08 | 5 | 43.60 ± 7.23 |
| 40 – 49 | 8 | 29.63 ± 11.50 | 51 | 46.69 ± 7.81 | 2 | 41.00 ± 1.41 |
| 50 – 59 | 6 | 29.83 ± 10.25 | 55 | 43.71 ± 10.65 | — | — ± — |
| 60 – 69 | 4 | 24.25 ± 6.70 | 39 | 38.15 ± 11.06 | 4 | 37.25 ± 6.90 |
| > 70 | 3 | 30.67 ± 5.03 | 61 | 25.79 ± 10.16 | — | — ± — |

*Note*. *N* = 522. South and A.R. = South and Autonomous Regions.

Finally, another two-way analysis of variance (age groups x place of residence) showed a non-significant effect on RSPM scores [$F_{(12, 476)}$ = 1.06; $p$ > .05]. The scores by age groups and regions are presented in Table 11.


DOI: 10.31211/rpics.2020.7.1.210

**Table 11**

*Descriptives of the Raven's Standard Progressive Matrices Scores by Age Groups x Place of Residence*

| Age groups | Urban | | Transition area | | Rural | |
|---|---|---|---|---|---|---|
| | *n* | *M* ± *SD* | *n* | *M* ± *SD* | *n* | *M* ± *SD* |
| 10 – 19 | 29 | 44.07 ± 7.98 | 2 | 48.50 ± 7.78 | 28 | 47.46 ± 6.25 |
| 20 – 29 | 52 | 47.33 ± 8.35 | 3 | 42.33 ± 6.11 | 14 | 46.21 ± 6.40 |
| 30 – 39 | 47 | 44.64 ± 9.82 | 3 | 38.33 ± 17.79 | 32 | 48.31 ± 6.89 |
| 40 – 49 | 49 | 43.80 ± 11.13 | 5 | 46.80 ± 17.31 | 25 | 43.80 ± 5.03 |
| 50 – 59 | 38 | 42.05 ± 12.21 | 3 | 34.67 ± 17.01 | 32 | 44.56 ± 7.80 |
| 60 – 69 | 22 | 36.82 ± 12.44 | 7 | 33.71 ± 10.75 | 33 | 35.45 ± 11.95 |
| > 70 | 20 | 28.45 ± 10.96 | 23 | 28.17 ± 11.83 | 30 | 24.90 ± 10.21 |

*Note*. *N* = 522.

## Psychometric properties

### *Reliability*

In the present study, a Cronbach's alpha of .94 was obtained for the total sample, indicating a high internal consistency among the RSPM items (Murphy & Davidshofer, 2004).

### *Correlations*

In Table 12, we can see correlations between Raven's Standard Progressive Matrices, Rey Complex Figure Test-Copy, Rey Complex Figure Test-Immediate Recall, and Rey Complex Figure Test-Delayed Recall were all high to moderate ($R^2$ between 22.1% and 77.4%). The correlation with Zung Self-Rating Anxiety Scale was low and negative ($R^2$ = 7.3%). Additionally, the correlational analysis showed no correlation between RSPM scores and the total time taken by the participants ($r$ = .09; $R^2$ = 8.1%; $p$ > .05).

**Table 12**

*Correlations between Raven's Standard Progressive Matrices and the other Study Measures*

| Measures | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1. RSPM | — | .55*** | .49*** | .47*** | -.27*** |
| 2. RCFT Copy | | — | .50*** | .48*** | -.27*** |
| 3. RCFT Immediate Recall | | | — | .88*** | -.22*** |
| 4. RCFT Delayed Recall | | | | — | -.24*** |
| 5. SAS Zung | | | | | — |

*Note*. *N* = 522. RSPM = Rey Complex Figure Test – Copy; RCFT = Rey Complex Figure Test; SAS Zung = Zung Self-Rating Anxiety Scale.
*** $p$ < .001.

### *Test-retest reliability*

To evaluate the temporal stability of RSPM, due to logistic reasons only a sub-sample of 12 participants was re-tested after the initial test administration (2.3%). The retest followed the same procedures as the initial testing. The test-retest interval averaged 4.67 months (*SD* = 1.37). Through the Student's *t* test for paired samples we found that there were no differences between the two moments [$t_{(11)} = 1.84$; $p > .05$; $d = 0.53$ (medium effect)], and the Pearson's correlation coefficient was .82 ($p < .001$), corresponding to a strong positive correlation between the two moments. The same was done for each sex, founding that there were no differences between the two moments, both for men [$t_{(3)} = 0.72$; $p > .05$; $d = 0.15$ (negligible effect)] and for women [$t_{(7)} = 1.65$; $p > .05$; $d = 0.32$ (small effect)] (Table 13).

**Table 13**

*Stability of Raven's Standard Progressive Matrices Scores Regarding the Global Sample and Sex Subgroups*

| | | *n* | *M* | *SD* | *r* |
|---|---|---|---|---|---|
| Total Group | Test | 12 | 41.67 | 11.68 | $.82^{***}$ |
| | Test-retest | 12 | 45.25 | 10.45 | |
| Men | Test | 4 | 46.75 | 13.60 | .94 |
| | Test-retest | 4 | 48.75 | 9.88 | |
| Women | Test | 8 | 39.13 | 10.63 | $.76^{*}$ |
| | Test-retest | 8 | 43.50 | 10.93 | |

*Note*. *N* = 522.
*** $p < .001$; * $p < .05$.

## Discussion

Our goal was to study the psychometric properties of Raven's Standard Progressive Matrices in a sample of the Portuguese population.

The total score of RSPM is 41.18, like others with similar samples (Al-Shahomee, 2012). According to findings in other normative studies, the performance of men is higher than women (Abdel-Khalek, 1988; Lynn & Irwin, 2004). Nevertheless, the effect size of the difference is small (also similar to Abdel-Khalek, 1988), which leads us to think that higher scores in men are not a universal phenomenon. Some authors state that men have an advantage regarding higher scores in general intelligence (Abdel-Khalek, 1988; Lynn & Irwin, 2004). Others claim that there are no significant differences between men and women (Colom & García-López, 2002; Court, 1983; Dutton et al., 2018; Jensen, 1998). According to the study of Carpenter et al. (1990), men get higher scores than women in certain types of items, namely, those that require a solution of addition/subtraction or a distribution of two rules. In other types of items, there are no sex differences in performance. Thus, the difference in our study on the

 DOI: 10.31211/rpics.2020.7.1.210

variable sex can relate to this question: different items on RSPM can measure different processes, and sexes differ on them (Mackintosh & Bennett, 2005). By contrast, Rushton et al. (2004) study shows that men have the same scores as women.

According to our study, there are higher scores in urban areas as in Al-Shahomee (2012). Adding to this, participants from the central region had higher scores. This region corresponds to Portugal's highest *per capita* incomes (PORDATA, 2015), and as is shown in the results and revision of the study of Almeida et al. (2011) and Brouwers et al. (2009), intelligence levels are correlated with *per capita* incomes.

Regarding levels of education, there are statistically significant differences between the different educational levels, as in Al-Shahomee (2012) and Brouwers et al. (2009). It was also possible to establish that a subject with a secondary or higher educational level gains a higher score than those who have only the basic education or have had no education. These data suggest that this test is not as insensitive to the acquired knowledge, scholastic or cultural content as Jensen pretended (1980). This result also suggests that intelligence assessment should not be based solely on nonverbal tests (Kaya et al., 2016).

In this study, a Cronbach's alpha was obtained for the total sample of .94, a value that must be considered quite satisfactory and is supported by Burke's study (2010). Other studies report that most of the coefficients exceed .90, with a modal value of .91 (review of Raven et al., 2000). Test-retest values from the original work of RSPM (between .83 and .93) reinforces the values found in our study (Raven et al., 2000). The Cronbach's alpha and test-retest are satisfactory, taking into account other studies (Abdel-Khalek, 1988; Moran, 1986). The correlation between scores on the two moments reinforces the good precision instrument temporal stability.

The correlational analysis of our study, which confirms the downward trend of the RSPM scores with age, was similar to the study of Burke (2010).

## Limitations and Conclusions

Our study also has some limitations that should be considered when interpreting the findings. First, the main limitation concerns temporal stability due to the small number of participants ($n$ = 12). For the paired samples $t$-test, a power of 0.95 and an alpha of .05, the sample would need 54 participants (Faul et al., 2007a, 2007b).

An additional limitation regards education level. The educational level was operationalized as the number of years of regular formal schooling completed with success. This approach is vulnerable to the numerous changes in the school system that has occurred throughout the last decades in Portugal. For instance, the *curricula* of regular school have suffered significant modifications. Since 2005, a large portion of the Portuguese adult population with low education has enrolled in the "New Opportunities" initiative to enhance their qualifications and to acquire primary (9th grade) or secondary (12th grade) level education certificates. However, the actual equivalence between regular school and "New Opportunities" education programs is still unknown. So, for the purpose of this study, only regular schooling was credited to the participants' educational background. Considering these continuing changes in education, future studies should assess the reliability of the normative algorithms and update the norms if necessary.

During the tests, the participants were not taking medication and had no symptoms of a disease that might potentially affect the performance of the tests. However, it is possible that there were participants suffering from undiagnosed conditions or mild cognitive deficit. In the future, more medical information should be gathered regarding the possible effect of disease/medication on test results.

Furthermore, more research is needed to develop studies of this nature, given the relative lack of published studies. Our aim in writing this article was to help overcome the lack of instruments adapted to the Portuguese population, and we hope our contribution will encourage further studies with the RSPM. The data from this study suggest that this is a tool with the potential to be used in the Portuguese population.

**Conflict of interest | Conflito de interesses**: None| Nenhum.

**Funding sources |Fontes de financiamento:** None| Nenhuma.

**Contributes: IG**: Literature review; Collection, data input and handling; Manuscript drafting. **HES**: Project coordination and design. Contribution to the writing of the manuscript; Data processing and discussion; Revision of the final draft of the manuscript. **CP**: Contribution to the writing of the manuscript; Collection, data input and handling; Revision of the final draft of the manuscript.

## References

Abdel-Khalek, A. M. (1988). Egyptian results on the Standard Progressive Matrices. *Personality and Individual Differences*, *9*(1), 193–195. https://doi.org/dfxbpd

Al-Shahomee, A. A. (2012). A standardisation of the Standard Progressive Matrices for adults in Libya. *Personality and Individual Differences*, *53*(2), 142–146. https://doi.org/gftztx

Alarcón Paz, C., Díaz Valenzuela, V., Hernández Rosales, J., & Estrada Goic, C. (2012). Estudio sobre la pertinencia del uso de las normas disponibles del Raven en adultos mayores chilenos [Study on the relevance of using available norms Raven in Chileans older adults]. *Psico-USF*, *17*(3), 387–395. https://doi.org/d22m

Almeida, L. S., Lemos, G., & Lynn, R. (2011). Regional Differences in Intelligence and per capita Incomes in Portugal. *Mankind Quarterly, 52*(2). https://doi.org/gf2x

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's Standard Progressive Matrices Test. *Assessment*, *19*(3), 354–369. https://doi.org/f34ssx

Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, *27*(4), 335–340. https://doi.org/fvc52t

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item recognition trial: A technique to enhance sensitivity of the Rey 15-Item memorization test. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, *24*(5), 561–573. https://doi.org/cdx69d

Brouwers, S. A., Van de Vijver, F. J. R., & Van Hemert, D. A. (2009). Variation in Raven’s Progressive Matrices scores across time and place. *Learning and Individual Differences, 19*(3), 330–338. https://doi.org/fwqjfd

Burke, H. R. (2010). Raven's Progressive Matrices: Validity, reliability, and norms. *The Journal of Psychology*, *82*(2), 253–257. https://doi.org/chrknh

Cairo Martínez, E., Bouza, C., Solozabal, T. P., & Cairo Valcárcel, E. (2000). Algunas caracteristicas y posibilidades del Test de Matrices Progresivas de Raven. *Revista Cubana de Psicologia*, *17*(2), 95–105. https://bit.ly/3mNBan7

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404–431. https://doi.org/fq7wtd

Carter, P. (2005). *The complete book of intelligence tests: 500 exercises to improve, upgrade and enhance your mind strength*. John Wiley and Sons Ltd.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. https://doi.org/fs6ptd

Chang, A. (2014). *StatsToDo* [Online calculator]. Brisbane, Queensland, Austrália. https://bit.ly/3tl76S5

Chiesi, F., Ciancaleoni, M., Galli, S., & Primi, C. (2012). Using the Advanced Progressive Matrices (Set I) to assess fluid ability in a short time frame: An item response theory–based analysis. *Psychological Assessment*, *24*(4), 892–900. https://doi.org/f4h27q

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Colom, R., & García-López, O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality and Individual Differences*, *32*(3), 445–451. https://doi.org/bzxjwp

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

Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). A Frontal Assessment Battery at bedside. *Neurology*, *55*(11), 1621–1626. https://doi.org/gd593r

Dutton, E., Bakhiet, S. F. A., Madison, G., Essa, Y. A. S., & Rajeh, M. Y. M. (2018). Sex differences on Raven's Standard Progressive Matrices within Saudi Arabia and across the Arab world: Females' advantage decreases from childhood to adolescence. *Personality and Individual Differences, 134*, 66–70. https://doi.org/gd7sm2

Espírito-Santo, H., Pires, C. F., Garcia, I. Q., Daniel, F., Silva, A. G. da, & Fazio, R. L. (2017). Preliminary validation of the Portuguese Edinburgh Handedness Inventory in an adult sample. *Applied Neuropsychology: Adult, 24*(3), 275–287. https://doi.org/ggqd8r

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007a). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/bxjdcg

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007b). *G*Power 3* (Version 3.1.9.2 for Mac OS X) [Computer Program]. Heinrich-Heine-Universität Dusseldorf. http://www.gpower.hhu.de/en.html

Garcia, I., Pessoa, I., Monteiro, B., Daniel, F., Lemos, L., & Espirito Santo, H. (2016). Propriedades psicométricas da versão Torga do Teste Stroop [Psychometric properties of Torga version of the Stroop Test]. *Revista Portuguesa de Investigação Comportamental e Social, 2*(2), 55–64. https://doi.org/gds5

Golan, D., Wilken, J., Doniger, G. M., Fratto, T., Kane, R., Srinivasan, J., Zarif, M., Bumstead, B., Buhse, M., Fafard, L., Topalli, I., & Gudesblatt, M. (2019). Validity of a multi-domain computerized cognitive assessment battery for patients with multiple sclerosis. *Multiple Sclerosis and Related Disorders, 30*, 154–162. https://doi.org/gf2z

Griffin, G. A. E., Glassmire, D. M., Henderson, E. A., & McCann, C. (1997). Rey II: Redesigning the Rey screening test of malingering. *Journal of Clinical Psychology*, *53*(7), 757–766. https://doi.org/cn88v3

Irrgang, M., Dorenkamp, M., Reohr, P., & Vik, P. (2019). C-28 Raven's Progressive Matrices: Validation of a short form. *Archives of Clinical Neuropsychology, 34*(6), 1057–1057. https://doi.org/f6tr

Irvine, S. H., & Berry, J. W. (Eds.). (1988). *Human abilities in cultural context*. Cambridge University Press. https://doi.org/cpjx8v

Jensen, A. R. (1980). *Bias in mental testing*. Free Press.

Jensen, A. R. (1987). Differential psychology: Towards consensus. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 413-467). The Falmer Press.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.

Kaplan, R. M., & Saccuzzo, D. P. (2012). *Psychological testing: Principles, applications, and issues* (8th ed.). Wadsworth, Cengage Learning.

Karnes, F. A., Lee, L. A., & May, B. (1982). Correlations among scores on the 1966, 1973 and 1979 norms of Raven's Standard Progressive Matrices for economically disadvantaged students. *Perceptual and Motor Skills*, *55*(3), 793–794. https://doi.org/dhvnrn

Kaya, F., Stough, L. M., & Juntune, J. (2017). Verbal and nonverbal intelligence scores within the context of poverty. *Gifted Education International, 33*(3), 257–272. https://doi.org/f6ts

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry e Endodontics*, *38*(1), 52–54. https://doi.org/gf8gxh

Kush, J. C. (1996). Fild-dependence, cognitive ability, and academic achievement in Anglo American and Mexican American students. *Journal of Cross-Cultural Psychology, 27*(5), 561–575. https://doi.org/crgwbg

Lovett, A., Forbus, K., & Usher, J. (2007). Analogy with qualitative spatial representations can simulate solving Raven's Progressive Matrices. In D. S. McNamara & J. C. Trafton (Eds), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 449–454). Cognitive Science Society.

Lynn, R., & Irwing, P. (2004). Sex differences on the progressive matrices: A meta-analysis. *Intelligence*, *32*(5), 481–498. https://doi.org/b3mtjt

Mackintosh, N. J., & Bennett, E. S. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. *Intelligence*, *33*(6), 663–674. https://doi.org/dvhwtx

Moran, A. P. (1986). The reliability and validity of Raven's Standard Progressive Matrices for Irish apprentices. *Applied Psychology*, *35*(4), 533–538. https://doi.org/dc9mvh

Murphy, K. R., & Davidshofer, C. O. (2004). *Psychological testing: Principles and applications* (6th ed). Prentice-Hall.

Orme, J. E. (1966). Hypothetically true norms for the Progressive Matrices Tests. *Human Development*, *9*(4), 222–229. https://doi.org/c3qtvf

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (6th ed.). McGraw Hill Education.

Pind, J., Gunnarsdóttir, E. K., & Jóhannesson, H. S. (2003). Raven's Standard Progressive Matrices: new school age norms and a study of the tests validity. *Personality and Individual Differences*, *34*(3), 375–386. https://doi.org/b9m7sq

PORDATA. (2015). Poder de compra per capita nos Municípios [Per capita purchasing power in the municipalities]. https://bit.ly/3sldvvo

Powers, S., & Barkan, J. H. (1986). Concurrent validity of the Standard Progressive Matrices for Hispanic and non-Hispanic and seventh-grade students. *Psychology in the Schools*, *23*(4), 333–336. https://doi.org/drkgx2

Raven, J. C. (1941). Standardization of Progressive Matrices, 1938. *British Journal of Medical Psychology*, *19*(1), 137–150. https://doi.org/dv6vk8

Raven, J., & Raven, J. (2003). Raven Progressive Matrices. In R. S. McCallum (Ed.), *Executive Functioning and Intelligence* (pp. 223–236). Springer US. https://doi.org/dcw6pd

Raven, J., Raven, J. C., & Court, J. H. (2000). *Manual for Raven's Standard Progressive Matrices and Vocabulary Scales. Section 3: Progressive Standard Matrices*. Harcourt Assessment.

Rey, A. (2002). Rey: *Teste de cópia de figuras complexas* (2nd ed.) [Rey: Copy complex figures test]. CEGOC-TEA.

Reznek, L. (2005). The Rey 15-item memory test for malingering: A meta-analysis. *Brain Injury*, *19*(7), 539–543. https://doi.org/dhht39

Rushton, J. P., Skuy, M., & Bons, T. A. (2004). Construct validity of Raven's Advanced Progressive Matrices for African and non-African engineering students in South Africa. *International Journal of Selection and Assessment*, *12*(3), 220–229. https://doi.org/dcfdt2

Savage-McGlynn, E. (2012). Sex differences in intelligence in younger and older participants of the Raven's Standard Progressive Matrices Plus. *Personality and Individual Differences*, *53*(2), 137–141. https://doi.org/d7hrcr

Serra, A. V., Ponciano, E., & Relvas, J. (1982). Aferição da escala de autoavaliação de ansiedade, de Zung, numa amostra de população portuguesa - II. - Sua avaliação como instrumento de medida [Self-evaluation Scale measurement of anxiety of Zung, a sample of the Portuguese population - II. — Its evaluation as an instrument of measurement]. *Psiquiatria Clínica*, *3*(4), 203–213.

Shulman, K. I., Shedletsky, R., & Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, *1*(2), 135–140. https://doi.org/b7vf79

Simões, M. M. R. (2000). *Investigações no âmbito da aferição nacional do Teste das Matrizes Progressivas Coloridas de Raven (M.P.C.R)* [Investigations under the national measurement test of Raven Coloured Progressive Matrices (RCPM)]. Fundação Calouste Gulbenkian.

Simões, M., Sousa, L., Duarte, P., Firmino, H., Pinho, M. S., Gaspar, N., Pires, L., Batista, P., Silva, A. R., Silva, S., Ferreira, A. R., & França, S. (2010). Avaliação da simulação ou esforço insuficiente com o Rey 15-Item Memory Test (15-IMT): Estudos de validação em grupos de adultos idosos [Evaluation of simulation or insufficient effort with Rey 15-Item Memory Test (15-IMT): Validation studies on groups of older adults]. *Análise Psicológica, 28*(1), 209–226. https://doi.org/gf22

Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, *19*(4), 465–473. https://doi.org/bx3kxr

Statistics Portugal. (2014). *Demographic statistics 2011*. INE.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/b77m95

Williams, J. E., & McCord, D. M. (2006). Equivalence of standard and computerized versions of the Raven Progressive Matrices Test. *Computers in Human Behavior*, *22*(5), 791–800. https://doi.org/b84zt7