Regression-Based Normative Data for Independent and Cognitively Active Spanish Older Adults: Free and Cued Selective Reminding Test, Rey–Osterrieth Complex Figure Test and Judgement of Line Orientation

Eva Calderón-Rubio 1, Javier Oltra-Cucarella 1,2,*, Beatriz Bonete-López 1,2, Clara Iñesta 1 and Esther Sitges-Maciá 1,2

Abstract: The aim of this work was to develop normative data for neuropsychological tests for the assessment of independent and cognitively active Spanish older adults over 55 years of age. Methods: regression-based normative data were calculated from a sample of 103 nondepressed independent community-dwelling adults aged 55 or older (66% women). The raw data for the Free and Cued Selective Reminding Test (FCSRT), the Rey–Osterrieth Complex Figure Test (ROCF) and the Judgement of Line Orientation Test (JLO) were regressed on age, sex and education. The model predicting the FCSRT delayed-recall (FCSRT-Del) scores also included the FCSRT immediate-recall (FCSRT-Imm) scores. The model predicting the ROCF immediate-recall (ROCF-Imm) scores included the ROCF copy-trial (ROCF-C) scores, and the model predicting the ROCF delayed-recall (ROCF-Del) scores included both the ROCF-C and the ROCF-Imm scores. In order to identify low scores, z-scores were used to determine the discrepancy between the observed and the predicted scores. The base rates of the low scores for both the SABIEX normative data and the published normative data obtained from the general population were compared. Results: the effects of the different sociodemographic variables (age, sex and education) varied throughout the neuropsychological measures. Despite finding similar proportions of low scores between the normative data sets, the agreement was irrelevant or only fair-to-good. Conclusions: the normative data obtained from the general population might not be sensitive enough to identify low scores in cognitively active older adults, incorrectly classifying them as cognitively normal compared to the less active population.

Keywords: mild cognitive impairment; memory; neuropsychological assessment; normative data

1. Introduction

The United Nations [1] predicts that, by 2050, the global population of older people will be growing at a rate of 2.6% per year. It is expected that 30% and 6% of the population in developed countries will be aged 60 or older and 80 or older, respectively. In Spain, the number of people aged 55 years or older is around 15 million, which represents 33.08% of the overall population. Amongst these, 18.22% (2,856,102) are 80 years or older [2].

The increase in the population’s age span is related to a raise in the probability of pathological ageing, such as cognitive impairment or dementia [3,4]. Indeed, age is the main risk factor for cognitive impairment [5,6] and Alzheimer’s disease (AD), with the prevalence of AD increasing from age 60 to 89 years [7]. This situation has an important impact on both public health and sanitary and social services [8]. On the one hand, cognitive impairment is a frequent cause of consultation in primary health care, with an estimated prevalence of 15–20% in individuals older than 60 years [9]. In Spain in particular, the
prevalence of cognitive impairment is 18.5% in people older than 65 years, and 45.3% in people older than 85 years [8]. On the other hand, the overall prevalence of dementia in the Spanish population is between 4% and 9% in groups older than 65 years [10]. Both cognitive impairment and dementia are related to higher comorbidity [11], susceptibility to infections [12] or depressive symptoms [13]. Furthermore, dementia has the highest total expenditure in neurological diseases [14] and requires about 3 h per day of care in both the basic and instrumental activities of daily living (ADL), becoming a huge burden for the caregivers of a family member with dementia [15].

A high percentage of individuals may not have a formal diagnosis of cognitive impairment or dementia: 55% in Spain [6], 58% in Europe [16] and 65% in the USA [17]. This situation could be due to the patients’ and relatives’ lack of consciousness about the presence of the disease [10], the lack of training among professionals in primary health care or the tight agendas in specialist settings [17]. For this reason, it is necessary to perform a thorough clinical study to overcome these problems. In order to detect subtle deficits or characterise cognitive strengths and weaknesses, an important part of this clinical study would be the neuropsychological assessment [18,19].

A neuropsychological assessment is based on the administration of standardised neurocognitive tests, in order to analyse the cognitive changes after brain damage [18,19]. It is essential to have reliable and suitable normative values, accounting for the effect of sociodemographic and cultural variables, in order to reduce the risk of misdiagnosis [20–22], determine the level of performance in the neuropsychological tests [23] and interpret the results obtained by a given subject by contrasting them with the performance of the reference group [18,24]. Normative data can be developed following three different strategies: mean and standard deviation scores, multiple regression and ROC curves. The first strategy is based on determining how the values are distributed and what their associated theoretical or empirical frequencies are [25–27]. The second is based on the generation of a prediction model accounting for age, sex and/or education. This model will be used to estimate a score, given certain variables (age, sex and education), that will be compared with the observed data [28–30]. ROC curves provide a cut-off value based on sensitivity and specificity [31].

Nevertheless, normative data obtained from the general population might not take into account the characteristics of active ageing. Active ageing is a process in which the opportunities for health, participation and security are optimised in order to enhance the quality of life during ageing [32]. This model encompasses six groups of determinants: behavioural styles, personal biological and psychological conditions, health and social services, physical engagement and social and economic factors. It has been previously reported that engaging in leisure activities is a relevant factor for a well-ageing process [33]. Indeed, most successfully ageing people, characterised by their health and independence, differ from the general population in the number of leisure activities they participate in [34]. Furthermore, cognitively stimulating activities may prevent cognitive decline during ageing [35].

Although educational level may play an important role in active ageing (for instance, [36]), the variable regarded as a determinant factor in the WHO model (WHO, 2002) is not early education, but lifelong learning [37]. In this sense, education is mostly taken into account as a long-term determinant [38]. Hijas-Gómez et al. [37] found that only the physical component was associated with survival and correlated with cognitive status, lifestyle and lifelong learning. For this reason, it is of great importance to ensure equal access opportunities for learning throughout the lifespan [39]. Thus, as active ageing is independent of the educational level, using normative data that account for educational level but do not comprehensively analyse the active ageing model could lead to a higher risk of misdiagnosis.

Two important cognitive domains assessed by neuropsychologists are memory and visuospatial perception [18,24]. In this work, memory was studied with the Free and Cued Selective Reminding Test (FCSRT; [40–42]) and the Rey–Osterrieth Complex Fig-
ure (ROCF) [43,44]. Visuospatial perception was analysed through the ROCF and the Judgement of Line Orientation (JLO; [45–47]) tests.

The FCSRT was initially designed as a selective reminding memory test [40,41], and then a cued-recall trial was added [42]. The ROCF measures the visual perception and visuospatial construction ability by means of the copy trial, and visual memory through the immediate- and delayed-recall tasks [43,44]. For both the FCSRT and ROCF, normative data in a healthy Spanish population have been developed by the NEURONORMA project [48,49]. The JLO analyses spatial perception and spatial orientation [46]. Normative data in a healthy Spanish population for JLO have also been developed by the NEURONORMA project [50].

Since cognitively active people seem to outperform nonactive people in neuropsychological tests, it is likely that the use of normative data that only account for educational level may increase the number of misdiagnoses [51]. For this reason, the objective of this work is to develop normative data for a cognitively active elderly population. We hypothesise that, accounting for the characteristics of the active ageing population, the low scores obtained with the normative data based on cognitively active people (SABIEX) will differ from those of the normative data obtained from the general population (NEURONORMA). Our hypothesis is that the active ageing population will show a higher percentage of low scores in the SABIEX normative data.

2. Materials and Methods

2.1. Participants

This was a cross-sectional observational study with cognitively healthy individuals living independently in the community. One hundred and five (105) students from the Seniors’ University (SABIEX) of the Universidad Miguel Hernández de Elche (UMH, Spain) were recruited from October 2019 to July 2021. SABIEX is an academic university programme for people aged 55 years or older covering subjects such as economics, psychology, politics and the arts.

The features of the sample have been previously described [52]. Inclusion criteria were (a) being 55 years old or older, (b) being cognitively normal (CN) without subjective cognitive complaints and (c) living independently in the community. Cognitive normality was determined by (a) Mini-Mental State Examination (MMSE; [53]) scores higher than 23, (b) Clinical Dementia Rating Scale (CDR; [54]) scores equal to 0 and (c) Lawton–Brody Instrumental Activities of Daily Living (IADL; [55]) scores equal to 7 or higher. Exclusion criteria were (a) unwillingness to participate in the neuropsychological assessment and (b) presence of vision and/or hearing impairments that might have interfered with the administration of cognitive tests. Having past or current medical conditions (e.g., cancer, psychiatric disorders or metabolic disease) was not defined as an exclusion criterion, so that the representativeness of the sample was guaranteed. All participants were born and raised in Spain and had Spanish as their mother tongue.

2.2. Materials

The individual neuropsychological assessment spanned around 90 min and covered different cognitive domains. In this paper, data are reported for tests assessing memory and visuospatial perception: the FCSRT [40,41], the ROCF [43,44] and the JLO [45–47] tests. Descriptive statistics for the remaining tests can be found in Bonete-López et al. [52].

2.2.1. Free and Cued Selective Reminding Test

The FCSRT is a word-list learning test administered through three immediate free and cued recall tasks and a 30 min delayed free and cued recall task. The FCSRT was administered according to NEURONORMA instructions [49]. Participants were shown 4 different sheets, all of them containing 4 bold written words. Each word pertained to a different semantic category. The four sheets were presented individually and in sequence. Subjects were asked to read each of the four words aloud and match each of them with
the semantic clue given by the examiner. After the 16 items were correctly identified, a nonsemantic interference task (counting backwards by three) was carried out for 20 s. After the interference task, the participants were asked to freely recall the 16 words in any order. Ninety seconds were given as the maximum time to complete the task. The semantic category was given (cued recall) for the words not recalled in the free-recall task. This task was followed by two more trials, each of them separated by a nonsemantic interference task (counting backwards by three). In the first two trials, participants were given the correct items for words not recalled with the semantic cue. After 30 min, the delayed-recall trial was performed. If any word was not remembered, the task was followed by the cued-recall trial. Lastly, 40 words were read out loud by the examiner [52]. Of them, 16 were words from the previous free- and cued-recall tasks (targets) and 24 were distractors. Subjects were asked to identify the targets by saying “yes” and the distractors by saying “no”. In this paper, the study variables were the total immediate- and total delayed-recall scores (FCSRT-Imm and FCSRT-Del, respectively) with a maximum score of 48 and 16, respectively.

2.2.2. Rey–Osterrieth Complex Figure

The ROCF test consisted of 18 elements, each of which could be scored 0.5, 1 or 2 points depending on their accuracy and location [24]. The ROCF test included three trials: (a) a copy of the figure model (ROCF-C), (b) an immediate-recall task (ROCF-Imm) and (c) a delayed-recall task after 30 min (ROCF-Del). For each trial, a blank sheet of paper was placed horizontally in front of the participants. In the first trial, participants were asked to copy the figure, trying not to leave any element undrawn and to respect the details and dimensions of the model. They were allowed to rotate their piece of paper, but not the model, which was just shown in the copy trial. For both the immediate- and delayed-recall trials, subjects were asked to draw as many items as they could remember. If they could not remember the location of the element, they were requested to draw it wherever they desired.

The study variables were: (a) copy—overall sum of correctly drawn elements when copying the model; (b) immediate recall—total sum of correctly drawn items after 3 min; and (d) delayed recall—total sum of correctly drawn items after 30 min. The maximum score was 36 for each task.

2.2.3. Judgement of Line Orientation

The JLO test consisted of five practice items and 30 test items. Each element was formed by two unnumbered segments and 11 numbered lines forming a semicircle as a model. The task was based on visually matching both test lines with two of the semicircle’s segments in each trial. The five practice tests were initially presented, and then the task continued with the performance of the test items. The total score was given by the number of elements in which both lines were correctly identified.

2.3. Procedure

The participants were invited to voluntarily participate in the study. No credit for their courses were given. The participants were individually assessed by a board-certified neuropsychologist and trained undergraduate, master’s degree-level or PhD-level students. An informed consent form was signed by every subject prior to enrolment. All participants provided personal and familial medical information. This project was approved by the UMH Ethics Committee (DPS.ESM.01.19).

2.4. Statistical Analysis

2.4.1. Calculation of Normative Data

We used the methodology reported by Iniesta et al. [56]. Briefly, a linear regression model was built to predict neuropsychological test scores using age, sex and education. To test for possible nonlinear associations, we centered age and education using the lowest value in the distribution of data, which are referred to as Age\textsuperscript{Min} and Education\textsuperscript{Min}.
throughout the manuscript, and calculated the quadratic term. We used a two-step linear regression model with raw values introduced in the first step and the quadratic terms introduced in the second step.

To improve the interpretation of performance on neuropsychological tasks conditional on other, previous neuropsychological tasks, we calculated normative data for the ROCF-Imm and ROCF-Del tests with a regression model adding ROCF-C scores to demographic variables, which are referred to as ROCF-Imm_{SABIEX} and ROCF-Del_{SABIEX}. In the regression model for the FCSRT-Del test, we added FCSRT-Imm scores to the demographic variables. Each variable was centered using the lowest value in the distribution of data.

2.4.2. Comparing Normative Data Sets

The number of low scores shown by our sample when using either the SABIEX or the NEURONORMA normative data was analyzed with the McNemar test (corrected for continuity) for related proportions [57], and the Fleiss’ kappa [57] interrater correlation coefficient for categorical data was used to analyse the level of agreement between normative data sets. According to Fleiss et al. [57], agreement beyond chance can be interpreted as poor, fair to good and excellent for values of 0–0.40, 0.41–0.75 and >0.75, respectively. Low scores were defined as scaled scores equal to or lower than 6 using NEURONORMA, and as z-scores equal to or lower than −1.28 using SABIEX normative data [56]. Using z-scores equal to or lower than −1.28 guarantees that the true positive rate and true negative rate from a linear regression are close to the 95% for a sample of size n = 100 or larger [58].

3. Results

From a pool of 105 subjects (33% males), two subjects were not included because of MMSE scores <24. Descriptive statistics for the demographic variables and the MMSE, IADL and GSD scores of the 103 participants are reported in Íñesta et al. [56]. A descriptive analysis of the complete test battery can be found in Bonete-López et al. [52].

The differences in age between the sexes were statistically significant (t (df = 101) = 3.06; p = 0.004). Specifically, the men were, on average, four years older than the women (men: M = 68.47, SD = 6.51; women: M = 64.42, SD = 6.22). No statistically significant differences between the sexes were found in the years of education (t (df = 101) = 0.551; p = 0.583), MMSE (t (df = 101) = −1.59; p = 0.114), IADL (t (df = 101) = 0.70; p = 0.485) and GDS (t (df = 101) = −1.18; p = 0.240).

The descriptive statistics for the performance in the different tasks of each neuropsychological test is provided in Table 1.

| Neuropsychological Test | M   | SD  | Range |
|-------------------------|-----|-----|-------|
| FCSRT-Imm               | 44.01| 4.04| 26–48 |
| FCSRT-Del               | 14.86| 1.48| 9–16  |
| ROCF-C                  | 28.18| 2.24| 14–36 |
| ROCF-Imm                | 14.97| 4.90| 2–26.5|
| ROCF-Del                | 15.04| 4.80| 2–25  |
| JLO                     | 20.55| 4.94| 9–30  |

M: mean; SD: standard deviation; FCSRT-Imm: Free and Cued Selective Reminding Test, total immediate recall; FCSRT-Del: Free and Cued Selective Reminding Test, total delayed recall; ROCF-C: Rey–Osterrieth Complex Figure, copy; ROCF-Imm: Rey–Osterrieth Complex Figure, immediate recall; ROCF-Del: Rey–Osterrieth Complex Figure, delayed recall; JLO: Judgement of Lines Orientation.

3.1. Calculation of Normative Data

The demographic variables (age, sex and education) showed different effects on each neuropsychological test. However, as no relation was found between the demographic variables and any of the tasks of the ROCF test, normative data were calculated with the means and standard deviations.
The multiple linear regression models of the rest of the neuropsychological measures are shown in Table 2. Sex was significantly related to the FCSRT-Imm, FCSRT-Del and JLO scores. Education$^{\text{Min}}$ was associated with the FCSRT-Imm, FCSRT-Del and FCSRT-Del$^{\text{SABIEX}}$ scores. Education$^{\text{Min}^2}$ had an effect on the FCSRT-Del and JLO scores. No statistically significant effects of Age$^{\text{Min}}$ and Age$^{\text{Min}^2}$ were found. Particularly interesting for this work are the significant associations within the dependent tasks, as the FCSRT-Del scores were related to the FCSRT-Imm$^{\text{Min}}$ scores, the ROCF-Imm scores to the ROCF-C$^{\text{Min}}$ scores and the ROCF-Del scores to the ROCF-Imm$^{\text{Min}}$ scores.

Table 2. Multiple linear regression models.

|               | B   | Std. Error | Sig. | R$^2$ | SEE  |
|---------------|-----|------------|------|-------|------|
| FCSRT-Imm     |     |            |      |       |      |
| Intercept     | 40.31 | 1.17       | 0.000 | 0.104 | 3.865 |
| Education$^{\text{Min}}$ | 0.288 | 0.111     | 0.011 |       |      |
| Sex           | 1.886 | 0.811      | 0.022 |       |      |
| FCSRT-Del     |     |            |      |       |      |
| Intercept     | 11.567 | 0.836     | 0.000 | 0.195 | 1.350 |
| Education$^{\text{Min}}$ | 0.547 | 0.185     | 0.004 |       |      |
| Sex           | 0.667 | 0.283      | 0.021 |       |      |
| FCSRT-Del$^{\text{SABIEX}}$ |     |            |      |       |      |
| Intercept     | 9.386 | 0.430      | 0.000 | 0.632 | 0.908 |
| Education$^{\text{Min}}$ | 0.272 | 0.023     | 0.000 |       |      |
| Education$^{\text{Min}^2}$ | −0.021 | 0.010 | 0.030 |       |      |
| ROCF-Imm$^{\text{SABIEX}}$ |     |            |      |       |      |
| Intercept     | 8.771 | 1.57       | 0.000 | 0.144 | 4.554 |
| ROCF-Del$^{\text{SABIEX}}$ |     |            |      |       |      |
| Intercept     | 3.727 | 0.616      | 0.000 | 0.792 | 2.199 |
| ROCF-C$^{\text{Min}}$ | 0.872 | 0.044     | 0.000 |       |      |
| JLO           |     |            |      |       |      |
| Intercept     | 22.252 | 0.958     | 0.000 | 0.208 | 4.446 |
| Sex           | −4.263 | 0.933     | 0.000 |       |      |
| Education$^{\text{Min}^2}$ | 0.014 | 0.007      | 0.038 |       |      |

FCSRT-Imm: Free and Cued Selective Reminding Test, total immediate recall; FCSRT-Del: Free and Cued Selective Reminding Test, total delayed recall, independently calculated from FCSRT-Imm; FCSRT-Del$^{\text{SABIEX}}$: Free and Cued Selective Reminding Test, total delayed recall, conditional on FCSRT-Imm; ROCF-Imm$^{\text{SABIEX}}$: Rey–Osterrieth Complex Figure, immediate recall, conditional on ROCF-C; ROCF-Del$^{\text{SABIEX}}$: Rey–Osterrieth Complex Figure, delayed recall, conditional on ROCF-C and ROCF-Imm; JLO: Judgement of Line Orientation.

3.2. Comparing Normative Data Sets

The frequency and cumulative percentage of the low scores for the three normative data sets (NEURONORMA, SABIEX$^{\text{INDEP}}$ and SABIEX$^{\text{DEP}}$) are provided in Table 3 and Figure 1.

Table 3. Frequency and accumulated percentage of NEURONORMA, SABIEX$^{\text{INDEP}}$ and SABIEX$^{\text{DEP}}$ low scores.

| Low Scores | NEURONORMA | SABIEX$^{\text{INDEP}}$ | SABIEX$^{\text{DEP}}$ |
|------------|------------|--------------------------|------------------------|
|            | Freq | Cum% | Freq | Cum% | Freq | Cum% |
| 0          | 71   | 100  | 66   | 100  | 57   | 100  |
| 1          | 23   | 31.1 | 17   | 35.9 | 31   | 44.7 |
| 2          | 3    | 8.8  | 12   | 19.5 | 12   | 14.6 |
| 3          | 4    | 5.9  | 5    | 7.8  | 2    | 2.9  |
| 4          | 1    | 2    | 2    | 2.9  | 1    | 1    |
| 5          | 1    | 1    | 1    | 1    | 0    | 0    |
| Total      | 103  |      | 103  |      | 103  |      |

Freq.: frequency of subjects showing the number of low scores in the Low Scores column. Cum%: cumulative percentage.
Int. J. Environ. Res. Public Health 2021, 18, x FOR PEER REVIEW 7 of 15

No statistically significant differences were found between the NEURONORMA and SABIEX\textsubscript{INDEP} data sets in the percentage of participants with one or more low scores ($\chi^2 (N = 103) = 0.842; p = 0.358$). The Fleiss’ kappa coefficient showed only a fair-to-good agreement for both data sets when identifying the participants showing one or more low scores ($\kappa = 0.586; p < 0.001$).

Conversely, statistically significant differences were found in the proportion of participants showing one or more low scores between the NEURONORMA and the SABIEX\textsubscript{DEP} data sets ($\chi^2 (N = 103) = 6.50; p = 0.011$). Furthermore, although the Fleiss’ kappa coefficient was statistically significant, the agreement between the normative data sets in identifying the low scores was only fair-to-good ($\kappa = 0.464; p < 0.001$).

3.3. Free and Cued Selective Reminding Test

Using the SABIEX normative data, 8.74% of the sample had at least one low score, whereas for the NEURONORMA data, 1.94% of the sample (two subjects) obtained one or more low scores. The McNemar test was statistically significant ($\chi^2 (N = 103) = 5.143; p = 0.023$), meaning that the proportions of low scores were not similar in both data sets. The Fleiss’ kappa coefficient showed a lack of agreement in identifying the low scores ($\kappa = 0.328; p = 0.001$) (see Table S1).

We also performed separate analyses of the proportion of low scores for the neuropsychological measures considered dependent from the others and compared them to the NEURONORMA data set. First, we compared both the FCSRT-Del and FCSRT-Del\textsubscript{SABIEX} scores with the data from NEURONORMA. In both the FCSRT-Del and FCSRT-Del\textsubscript{SABIEX} data, 11.65% of the sample (12 subjects) had at least one low score, in contrast to the 0.97% (one subject) from NEURONORMA. The McNemar test showed statistically significant differences in the proportion of both the SABIEX and NEURONORMA data sets ($\chi^2 (N = 103) = 9.091; p = 0.002$) and no agreement was found in identifying the low scores ($\kappa = 0.097; p = 0.326$) (see Tables S2 and S3).

3.4. Rey–Osterrieth Complex Figure

The ROCF-C scores were not associated statistically with any of the variables analysed (see Table 2). Regarding the ROCF-Imm and ROCF-Imm\textsubscript{SABIEX} scores, at least one low score was obtained by 10.68% of the sample (11 subjects) in the SABIEX\textsubscript{INDEP}, 8.74% (9 subjects) in the SABIEX\textsubscript{DEP} and 6.80% (7 subjects) in the NEURONORMA data set. The McNemar test showed no statistically significant differences when comparing the
NEURONORMA data with either the SABIEX\textsubscript{INDEP} ($\chi^2 (N = 103) = 1.500; p = 0.221$) or the SABIEX\textsubscript{DEP} ($\chi^2 (N = 103) = 0.125; p = 0.7237$) data. The Fleiss’ kappa coefficient showed fair-to-good agreement between the NEURONORMA data and both of the SABIEX data sets (SABIEX\textsubscript{INDEP}: $\kappa = 0.635; p = 0.000$; SABIEX\textsubscript{DEP}: $\kappa = 0.458; p < 0.001$) (see Tables S4 and S5).

Regarding the ROCF-Del and ROCF-Del\textsubscript{SABIEX} scores, at least one low score was obtained by 12.62% of the sample (13 subjects) in the SABIEX\textsubscript{INDEP}, 10.68% (11 subjects) in the SABIEX\textsubscript{DEP} and 9.71% (10 subjects) in the NEURONORMA data set. The McNemar test showed no statistically significant differences when comparing the NEURONORMA data with either the SABIEX\textsubscript{INDEP} ($\chi^2 (N = 103) = 0.800; p = 0.371$) or the SABIEX\textsubscript{DEP} ($\chi^2 (N = 103) = 0.000; p > 0.999$) data sets. However, the Fleiss’ kappa coefficient was only statistically significant when comparing the NEURONORMA data with the SABIEX\textsubscript{INDEP} data ($\kappa = 0.755; p = 0.000$), with a lack of agreement in identifying the participants as showing one or more low scores between the SABIEX\textsubscript{DEP} and NEURONORMA data sets ($\kappa = -0.114; p = 0.249$) (see Tables S6 and S7).

3.5. Judgement of Line Orientation

The JLO scores were negatively associated with the sex variable and positively associated with the Education\textsuperscript{MIN2} variable, meaning that the effect of education was not linear and that better scores are expected when the level of education increases.

4. Discussion

The objective of this study was to develop normative data for Spanish cognitively active individuals aged 55 or older. For this aim, regression analyses were performed for every variable studied. Two novelties characterise this work: the use of a cognitively active sample and the analysis of every subtest as both dependent (SABIEX\textsubscript{DEP}) and independent (SABIEX\textsubscript{INDEP}) of other variables from the same test. Our results showed only a fair-to-good agreement in identifying the low scores between both the SABIEX normative data and the NEURONORMA normative data, with no agreement for the delayed-recall trial from the ROCF test.

No effect of age was found in any of the tests assessed, which differs from previous studies suggesting that neuropsychological performance declines with advancing age [49,59–62]. This result may represent a novelty in this area, as it may show that age becomes less relevant when individuals are cognitively active. This may also evidence the importance of creating a cognitive reserve [63] throughout the lifespan, independently of the level of education achieved during the early life stages.

Regarding the FCSRT, sex had a significant effect on both the FCSRT-Imm and FCSRT-Del trials, in line with previous studies suggesting that women outperform men on tests based on verbal material such as the FCSRT [64–66]. Other studies also showed a significant effect of sex in the free-recall tasks of the FCSRT [67,68]. However, its effect may be minor or irrelevant [49,69,70] and may be influenced by the different proportion of men–women in our study. The effect of sex became non-significant when we analysed the FCSRT-Del scores while controlling for the FCSRT-Imm scores (FCSRT-Del\textsubscript{SABIEX}). Education also showed a significant effect on the FCSRT-Imm, FCSRT-Del and FCSRT-Del\textsubscript{SABIEX} scores, in line with previous studies [49,66–70]. This result may illustrate the effect of education on neuropsychological performances and its association with cognitive reserve [61,69–71]. Quadratic education had an impact only on the FCSRT-Del scores, meaning that the effect of education was not linear and that the differences in memory performance were more significant amongst less educated subjects than between highly educated individuals, which is in line with previous works [72–75]. This effect disappeared when the FCSRT-Imm score was also controlled.

Regarding the ROCF test, no significant effect of any of the sociodemographic variables was found. In previous studies, minor or nonexistent effects have been reported [49,76–78], which is in line with our results. However, most of the scientific research shows that education has a positive impact when performing ROCF tasks, as higher education levels
are usually associated with better scores [49,78–80]. Nevertheless, our results may indicate that, independently of the level of education, when a cognitively active lifestyle is carried out, cognitive strategies can compensate for these differences [81], as suggested by the cognitive reserve construct [63,82]. Lastly, regarding the JLO test, sex and quadratic education showed a significant effect, which is in line with previous studies [83–85].

Another relevant aspect of our research was the analysis of the different subtests as both independent of and dependent on the rest of the trials of the same test. In line with previous results [78], the ROCF-C scores had a positive effect on the ROCF-Imm scores and the ROCF-Imm on the ROCF-Del, and the FCSR-T-Imm scores had a significant positive effect on the FCSR-T-Del scores. When comparing the SABIEXINDEP and NEURONORMA data sets, we found similar proportions of low scores, but there was only a fair-to-good agreement in identifying the low scores. On the other hand, when comparing the SABIEXDEP and NEURONORMA data sets, we found a statistically different proportion of low scores and only a fair-to-good agreement, which was even lower than that of the SABIEXINDEP data set. This may have been due to the distinctive features characterising our sample: the NEURONORMA data was based on the general population, while the SABIEX sample included highly cognitively active participants. This may suggest that using normative data that do not take into account the specificities of the active ageing population may increase the number of diagnosis errors and, hence, the misclassification of subjects as cognitively impaired if low scores are to be used to diagnose cognitive impairment. Additionally, not controlling the effect of the related subtests when creating new normative data sets, as is the usual practice [49,79,80,84], may also be associated with an increased rate of false positives and/or negatives and with the misclassification of subjects. Further research is needed to confirm the clinical applicability of our results, analysing whether the use of SABIEX normative data is useful to identify with greater certainty individuals with a greater risk of cognitive decline. Lastly, when contrasting the neuropsychological measures separately, no agreement (FCSR-T-Imm, FCSR-T-Del, FCSR-T-DelSABIEX, ROCF-DelSABIEX) or only a fair-to-good agreement (ROCF-Imm, ROCF-ImmSABIEX, ROCF-Del) was found in identifying low scores. This supports the idea that normative data accounting for the characteristics of cognitively active elders are needed to avoid the appearance of false negatives and/or false positives and, consequently, their misclassification.

These results may have important clinical implications. For example, objective cognitive impairment (i.e., one or more low scores) is necessary for the diagnosis of MCI [86,87]. Our results are important whether classic [86,88] or modified [89–91] diagnostic criteria are to be used to identify MCI. If only one test is used to identify memory impairments, as with the classic criteria, the selection of the appropriate normative sample against which the raw scores are to be compared may reduce the number of false positives. If several tests are used to identify MCI, as with the modified diagnostic criteria, the association between several tasks within the same test might identify with greater certainty individuals with true cognitive impairment. For example, if two low scores on the memory task are needed to meet the criterion for the comprehensive diagnostic criteria [89,92], using the delayed-recall scores as independent of the immediate-recall scores might be associated with a higher number of false positives, because the correlation between the tests is not taken into account when independent normative data are used. Relatedly, analysing the performance on some tests that are conditional on other tests might impact the number of low scores needed to define normal variability and, subsequently, true cognitive impairment [90,91]. Using the worst-performing 10% of the sample, two or more low scores would be needed to meet the criterion for objective cognitive impairment using the NEURONORMA normative data, compared to three or more low scores using the SABIEX normative data (Table 3). Additionally, since low scores on visual memory tasks have been associated with a similar risk of progression from MCI to AD compared to low scores on verbal memory tasks [93], using the normative data for the delayed-recall trial in the ROCF test conditional on the copy trial may help clinicians to better differentiate true impairments in visual memory from low scores on the immediate or delayed trials that are conditional on low scores on the copy trial.
If low scores on the delayed-recall task are associated with low scores on the copy trial, the individual would be diagnosed as having nonamnestic MCI rather than as having amnestic MCI, with the former being associated with a lower risk of progression from MCI to AD than the latter [94]. We provide a friendly calculator for researchers and clinicians that is available at https://docs.google.com/spreadsheets/d/1p-RDT6F85EsXPxALV-l6R8Sq-E4qGEr0/edit?usp=sharing&ouid=109893127231470805500&rtpof=true&sd=true (accessed on 8 December 2021).

This work has certain limitations. First, we used an incidental sampling method, which was related to the overrepresentation of people aged 60–70 and the underrepresentation of people aged over 75 years. It was also associated with the overrepresentation of women. As gender differences in cognitive domains have been previously suggested [95,96], these results should be interpreted with caution. The number of younger and older subjects, as well as the number of men, may be further investigated to improve the reliability and accuracy of the data set. Secondly, with our methodology, we cannot determine the clinical applicability of the SABIEX normative data. We can only suggest that individuals showing low scores might differ between the SABIEX and population-based normative data sets, but further research into the clinical population is needed to find out which is more accurate in identifying cognitive impairment in cognitively active older adults if low scores are to be used to identify cognitive impairment. Lastly, this work was only based on the Spanish population, so, due to the cultural differences in neuropsychological functions suggested by other authors [21,22,97], interpretations and application in other cultures or languages should be taken with caution.

5. Conclusions

Our results suggest that normative data obtained from the general population might not be sensitive enough to identify low scores in cognitively active older adults. Normative data accounting for the characteristics of cognitively active older adults might be necessary to reduce the number of diagnostic errors and, eventually, their misclassification as cognitively impaired compared to normative data obtained from the general and less cognitively active population.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijerph182412977/s1, Table S1: Comparing number of low scores on FCSRT-Imm between normative data sets (NEURONORMA and SABIEX), Table S2: Comparing number of low scores on FCSRT-Del between normative data sets (NEURONORMA and SABIEX_INDEP), Table S3: Comparing number of low scores on FCSRT-Del independent (NEURONORMA) with FCSRT-Del conditional on FCSRT-Imm (FCSRT-Del_SABIEX), Table S4: Comparing number of low scores on ROCF-Imm between normative data sets (NEURONORMA and SABIEX_INDEP), Table S5: Comparing number of low scores on ROCF-Imm independent (NEURONORMA) with ROCF-Imm conditional on ROCF-C (ROCF-Imm_SABIEX), Table S6: Comparing number of low scores on ROCF-Del between normative data sets (NEURONORMA and SABIEX_INDEP), Table S7: Comparing number of low scores on ROCF-Del independent (NEURONORMA) with ROCF-Del conditional on ROCF-Imm (ROCF-Del_SABIEX).

Author Contributions: Conceptualization, J.O.-C., E.S.-M. and B.B.-L.; methodology, J.O.-C., E.C.-R. and C.I.; formal analysis, E.C.-R. and J.O.-C.; investigation, E.C.-R., J.O.-C., B.B.-L., C.I. and E.S.-M.; data curation, E.C.-R. and J.O.-C.; writing—original draft preparation, E.C.-R. and J.O.-C.; writing—review and editing, E.C.-R., J.O.-C., B.B.-L., C.I. and E.S.-M.; supervision, J.O.-C., B.B.-L. and E.S.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially funded by the Conselleria d’Innovació, Universitats, Ciència i Societat Digital, Generalitat Valenciana (NEUROPREVENT project, GV/2021/139).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the Miguel Hernández University (protocol code DPS.ESM.01.19).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.
Data Availability Statement: Data are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations. World Population Ageing 2009; United Nations: New York, NY, USA, 2009.
2. Instituto Nacional de Estadística Proporción de Personas Mayores de Cierta Edad Por Comunidad Autónoma. Available online: https://www.ine.es/jaxiT3/Datos.htm?i=1451#tabs-tabla (accessed on 13 March 2021).
3. De Pedro-Cuesta, J.; Virués-Ortega, J.; Vega, S.; Seijo-Martínez, M.; Saz, P.; Rodríguez, F.; Rodríguez-Laso, A.; Reñé, R.; de las Heras, S.P.; Mateos, R.; et al. Prevalence of Dementia and Major Dementia Subtypes in Spanish Populations: A Reanalysis of Dementia Prevalence Surveys, 1990–2008. BMC Neuro. 2009, 9, 55. [CrossRef] [PubMed]
4. De Ronchi, D.; Berardi, D.; Menchetti, M.; Ferrari, G.; Serretti, A.; Dal Monte, E.; Fratiglioni, L. Occurrence of Cognitive Impairment and Dementia after the Age of 60: A Population-Based Study from Northern Italy. Dement. Geriatr. Cogn. Disord. 2005, 19, 97–105. [CrossRef]
5. Rodríguez-Sánchez, E.; Mora-Simón, S.; Patino-Alonso, M.C.; García-García, R.; Escribano-Hernández, A.; García-Ortiz, L.; Pere-Bartolomé, M.V.; Gómez-Marcos, M.A. Prevalence of Cognitive Impairment in Individuals Aged over 65 in an Urban Area: DERIVA Study. BMC Neuro. 2011, 11, 147. [CrossRef]
6. Tola-Arribas, M.A.; Yugueros, M.I.; Garea, M.J.; Ortega-Valín, F.; Cerón-Fernández, A.; Fernández-Malvido, B.; San José-Gallegos, A.; González-Touya, M.; Botrán-Velicia, A.; Iglesias-Rodriguez, V.; et al. Prevalence of Dementia and Subtypes in Valladolid, Northwestern Spain: The DEMINVALL Study. PLoS ONE 2013, 8, e77668. [CrossRef] [PubMed]
7. Jellinger, K.A.; Attems, J. Prevalence of Dementia Disorders in the Oldest-Old: An Autopsy Study. Neurology 2018, 83, 491–498. [CrossRef]
8. Petersen, R.C. Mild Cognitive Impairment. Contin. Minneap. Minn. 2016, 22, 404–418. [CrossRef]
9. Villarejo Galende, A.; Eimil Ortiz, M.; Llamas Velasco, S.; Llanero Luque, M.; López de Silanes de Miguel, C.; Prieto Jurczynska, C. Informe de la Fundación del Cerebro. Impacto social de la enfermedad de Alzheimer y otras demencias. Neurologia 2021, 36, 39–49. [CrossRef]
10. Poblador-Plou, B.; Calderón-Larrañaga, A.; Marta-Morenos, J.; Hanco-Saavedra, J.; Sicras-Mainar, A.; Soljak, M.; Prados-Torres, A. Comorbidity of Dementia: A Cross-Sectional Study of Primary Care Older Patients. BMC Psychiatry 2014, 14, 84. [CrossRef] [PubMed]
11. Matusik, P.; Tomaszewski, K.; Chmielowska, K.; Nowak, J.; Nowak, W.; Parnicka, A.; Dubiel, M.; Gałkowski, J.; Grodzicki, T. Severe Frailty and Cognitive Impairment Are Related to Higher Mortality in 12-Month Follow-up of Nursing Home Residents. Arch. Gerontol. Geriatr. 2012, 55, 22–24. [CrossRef]
12. Leyhe, T.; Reynolds, C.F.; Melcher, T.; Limmennan, C.; Klöppel, S.; Blennow, K.; Zetterberg, H.; Dubois, B.; Lista, S.; Hampel, H. A Common Challenge in Older Adults: Classification, Overlap, and Therapy of Depression and Dementia. Alzheimer’s Dement. 2017, 13, 59–71. [CrossRef]
13. Gustavsson, A.; Svensson, M.; Jacob, F.; Allgulander, C.; Alonso, J.; Beghi, E.; Dodel, R.; Ekman, M.; Faravelli, C.; Fratiglioni, L.; et al. Cost of Disorders of the Brain in Europe 2010. Eur. Neuropsychopharmacol. 2011, 21, 718–779. [CrossRef]
14. Lin, C.-Y.; Shih, P.-Y.; Ku, L.-J.E. Activities of Daily Living Function and Neuropsychiatric Symptoms of People with Dementia and Caregiver Burden: The Mediating Role of Caregiving Hours. Arch. Gerontol. Geriatr. 2019, 81, 25–30. [CrossRef]
15. Eefsting, J.A.; Boersma, F.; Brink, W.V.D.; Tilburg, W.V. Differences in Prevalence of Dementia Based on Community Survey and General Practitioner Recognition. Psychol. Med. 1996, 26, 1223–1230. [CrossRef]
16. Valcour, V.G.; Masaki, K.H.; Curb, J.D.; Blanchette, P.L. The Detection of Dementia in the Primary Care Setting. Arch. Intern. Med. 2000, 160, 2964–2968. [CrossRef] [PubMed]
17. Muro, M.D.; Howieson, D.B.; Bigler, E.D.; Tranel, D. Neuropsychological Assessment, 5th ed.; Oxford University Press: Oxford, UK; New York, NY, USA, 2012; ISBN 978-0-19-539552-5.
18. Munro Collum, C.; Lacritz, L. Neuropsychological Assessment in Dementia. In The American Psychiatric Publishing Textbook of Alzheimer Disease and Other Dementias; American Psychiatric Publishing: Washington, DC, USA, 2009.
19. Arango-Lasprilla, J.C.; Rivera, D.; Ramos-Usuga, D.; Vergara-Moragues, E.; Montero-López, E.; Adana Díaz, L.A.; Aguayo Arellis, A.; García-Guerrero, C.E.; García De La Cadena, C.; Llerena Espezúa, X.; et al. Trail Making Test: Normative Data for the Latin American Spanish-Speaking Pediatric Population. NeuroRehabilitation 2017, 41, 627–637. [CrossRef]
20. Ardila, A. Cultural Values Underlying Psychometric Cognitive Testing. Neuropsychol. Rev. 2005, 15, 185. [CrossRef] [PubMed]
21. Ardila, A. The Impact of Culture on Neuropsychological Test Performance. In International Handbook of Cross-Cultural Neuropsychology; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2007; pp. 23–44.
22. Horton, A.M.; Roberts, D. Demographic Effects on the Trail Making Test in a Drug Abuse Treatment Sample. Arch. Clin. Neuropsychol. Off. J. Natl. Acad. Neuropsychol. 2003, 18, 49–56. [CrossRef]
23. Strauss, E.; Sherman, E.M.S.; Spreen, O.; Spreen, O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary, 3rd ed.; Oxford University Press: Oxford, UK; New York, NY, USA, 2006; ISBN 978-0-19-515957-8.
25. Hall, J.R.; Balldin, V.H.; Gamebo, A.; Edwards, M.L.; Johnson, L.A.; O’Bryant, S.E. Texas Mexican American Adult Normative Studies: Neurological Data for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). Dent. Neuropsychol. 2018, 43, 27–35. [CrossRef]

26. O’Bryant, S.E.; Edwards, M.; Johnson, L.; Hall, J.; Gamebo, A.; O’jile, J. Texas Mexican American Adult Normative Studies: Neurological Data for Commonly Used Clinical Neuropsychological Measures for English- and Spanish-Speakers. Dent. Neuropsychol. 2018, 43, 1–26. [CrossRef]

27. Silvestre, G.; Iglesias, R.M.; Silvestre, E. Boston Naming Test Norms for the Dominican Population. Aphasiology 2018, 12, 340–365. [CrossRef]

28. Heaton, A.; Gooding, A.; Chernert, M.; Umlauf, A.; Franklin, D.R.; Rivera Minndt, M.; Suarez, P.; Artiola i Fortuni, L.; Heaton, R.K.; Marquione, M. Demographically-Adjusted Norms for the Grooved Pegboard and Finger Tapping Tests in Spanish-Speaking Adults: Results from the Neuropsychological Norms for the U.S.-Mexico Border Region in Spanish (NP-NUMBRS) Project. Clin. Neuropsychol. 2021, 35, 419–432. [CrossRef]

29. Peña-Casanova, J.; Blesa, R.; Aguilar, M.; Gramunt-Fombuena, N.; Gómez-Ansón, B.; Oliva, R.; Molinuevo, J.L.; Robles, A.; Barquero, M.S.; Antúnez, C.; et al. Spanish Multicenter Normative Studies (NEURONORMA Project): Methods and Sample Characteristics. Arch. Clin. Neuropsychol. 2009, 24, 307–319. [CrossRef]

30. Rivera, D.; Olabarrieta-Landa, L.; Van Der Elst, W.; Gonzalez, I.; Rodriguez-Aguado, Y.; Arelis, A.A.; Rodriguez-Irizarry, W.; De La Cadena, C.G.; Arango-Lasprilla, J.C. Normative Data for Verbal Fluency in Healthy Latin American Adults: Letter M, and Fruits and Occupations Categories. Neurologia 2019, 33, 287–300. [CrossRef]

31. Mandrekar, J.N. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. J. Thorac. Oncol. 2010, 5, 1315–1316. [CrossRef] [PubMed]

32. WHO—World Health Organization. Active Ageing: A Policy Framework; WHO: Madrid, Spain, 2002.

33. Lee, P.-L.; Lan, W.; Yen, T.-W. Aging Successfully: A Four-Factor Model. Educ. Gerontol. 2011, 37, 210–227. [CrossRef]

34. Fernández-Ballesteros, R.; Robine, J.M.; Walker, A.; Kalache, A. Active Aging: A Global Goal. Curr. Gerontol. Geriatr. Res. 2013, 1–4. [CrossRef] [PubMed]

35. Olivera, J.; Andreoli, F.; Leist, A.K.; Chauvel, L. Inequality in Old Age Cognition across the World. Arch. Gerontol. Geriatr. 2018, 25, 907–910. [CrossRef]

36. Benton, A.; Sivan, A.B.; Hamsher, K.D.; Varney, N.R.; Spreen, O. Contributions to Neuropsychological Assessment: A Clinical Manual; Oxford University Press: Oxford, UK, 1994.

37. Benton, A.L.; Varney, N.R.; Hamsher, K.D. Visuospatial Judgment. A Clinical Test. Arch. Neurol. 1978, 35, 364–367. [CrossRef]

38. Palomo, R.; Casals-Coll, M.; Sánchez-Benavidez, G.; Quintana, M.; Manero, R.M.; Rognoni, T.; Calvo, L.; Aranciva, F.; Tamayo, F.; Peña-Casanova, J. Spanish Normative Studies in Young Adults (NEURONORMA Young Adults Project): Norms for the Rey-Osterrieth Complex Figure (Copy and Memory) and Free and Cued Selective Reminding Test. Neurologia 2013, 28, 226–235. [CrossRef]

39. Osterrieth, P.A. Le Test de Copie d’une Figure Complexe; Contribution à l’étude de La Perception et de La Mémoire. Arch. Psychol. 1941, 30, 206–356.

40. Rey, A. L’examen Psychologique Dans Les Cas d’encéphalopathie Traumatique. (Les Problems.). Arch. Psychol. 1941, 28, 215–285.

41. Benton, A.; Hannay, H.J.; Varney, N.R. Visual Perception of Line Direction in Patients with Unilateral Brain Disease. Neurology 1973, 25, 907–910. [CrossRef]

42. Buschke, H. Fuld, P.A. Evaluating Storage, Retention, and Retrieval in Disordered Memory and Learning. Neurology 1974, 24, 1019–1025. [CrossRef] [PubMed]

43. Rey, A. L’examen Psychologique Dans Les Cas d’encéphalopathie Traumatique. (Les Problems.). Arch. Psychol. 1941, 28, 215–285.

44. Benton, A.; Sivan, A.B.; Hamsher, K.; Varney, N.R.; Spreen, O. Contributions to Neuropsychological Assessment: A Clinical Manual; Oxford University Press: Oxford, UK, 1994.

45. Benton, A.L.; Varney, N.R.; Hamsher, K.D. Visuospatial Judgment. A Clinical Test. Arch. Neurol. 1978, 35, 364–367. [CrossRef]

46. Palomo, R.; Casals-Coll, M.; Sanchez-Benavides, G.; Quintana, M.; Manero, R.M.; Rognoni, T.; Calvo, L.; Aranciva, F.; Tamayo, F.; Pena-Casanova, J. Spanish Normative Studies in Young Adults (NEURONORMA Young Adults Project): Norms for the Rey-Osterrieth Complex Figure (Copy and Memory) and Free and Cued Selective Reminding Test. Neurologia 2013, 28, 226–235. [CrossRef]
