

1 Petronilla Battista^{†*1,2}, Chiara Griseta^{†3}, Rosanna Tortelli^{4,5*}, Pietro Guida², Fabio Castellana³, Davide
2 Rivolta⁶, Giancarlo Logroscino^{4,7}.

3 *† These authors contributed equally as first authors of this study.*

4

5 **The Modified Five-Point Test (MFPT): Normative data for a sample of Italian Elderly**

6

7 1 Global Brain Health Institute, University of California, San Francisco, 675 Nelson Rising Lane, Suite 190,
8 San Francisco, CA 94518, USA

9 2 Istituti Clinici Scientifici Maugeri IRCCS, Pavia, Italy

10 3 Frailty Phenotypes Research Unit, "Salus in Apulia Study", National Institute of Gastroenterology "Saverio
11 de Bellis", Research Hospital, Castellana Grotte, Italy.

12 4 Center for Neurodegenerative Diseases and the Aging Brain, Department of Clinical Research in
13 Neurology, University of Bari Aldo Moro, Pia Fondazione Cardinale G Panico, Tricase, Italy

14 5 Institute of Neurology, University College London (UCL), London, UK

15 6 Department of Education, Psychology and Communication, University of Bari Aldo Moro, Bari, Italy

16 7 Department of Basic Medical Sciences, Neuroscience and Sense Organs, University of Bari Aldo Moro,
17 Bari, Italy

18

19 ***Corresponding authors:**

20 Dr Petronilla Battista, Global Brain Health Institute, University of California, San Francisco, USA

21 Email: petronilla.battista@gbhi.org

22 &

23 Dr Rosanna Tortelli, Center for Neurodegenerative Diseases and the Aging Brain, Department of Clinical
24 Research in Neurology, University of Bari Aldo Moro, Pia Fondazione Cardinale G Panico, Tricase, Italy

25 Email: r.tortelli@ucl.ac.uk

26

27 ORCID

28 Petronilla Battista 0000-0002-3120-1214

29

1 **Acknowledgements:** We wish to thank Alessandra Grasso, Cristina Didio, Annalaura Lagioia, Simona
2 Spaccavento, for their helpful support with data collection and/or scoring. This manuscript is also the result
3 of the research work on frailty undertaken by the “Italia Longeva: Research Network on Aging” team,
4 supported by the resources of the Italian Ministry of Health—Research Networks of National Health
5 Institutes. We thank the General Practitioners of Castellana Grotte, for the fundamental role in the
6 recruitment of participants to this studies: Campanella Cecilia Olga Maria, Daddabbo Annamaria, Dell’Era
7 Giosue’, Giustiniano Rosalia Francesca, Guzzoni Iudice Massimo, Lomuscio Savino, Lucarelli Rocco,
8 Mazzarisi Antonio, Palumbo Mariana, Persio Maria Teresa, Pesce Rosa Vincenza, Puzzovivo Gabriella,
9 Romano Pasqua Maria, Sgobba Cinzia, Simeone Francesco, Tartaglia Paola, Tauro Nicola.

10

11 **Funding:** This study was funded by the Italian Ministry of Health with “Ricerca Corrente 2019 Grant”.

12

13 **Author Contributions** PB designed the study with the input of RT and GL. CG and FC were involved in
14 participant recruitment. Eligibility, clinical examinations and collection were performed by PB, CG, RT and
15 FC processed and analysed the subject’s data. PG performed the statistical analysis; PB, CG, RT and DR
16 interpreted the data and wrote the manuscript; and all authors contributed to reviewing the manuscript.

17

18 **Compliance with ethical standards**

19 **Conflict of interests:** none

20 **Availability of data:** data are available upon request.

21

22

1 **Abstract**

2 Non-verbal figural fluency is related to executive functions and specifically to the ability to create as many unique
3 designs as possible, while minimizing their repetitions. An Italian version of figural fluency is the Modified Five Point
4 Test (MFPT), which is highly employed in the clinical practice of neuropsychologists. To date, reference data of Italian
5 population are limited to a sample aged between 16 and 60 years old. Thus, the current study aims to provide normative
6 data of the MFPT in the context of a population-based setting, conducted in Southern Italy, aged over 65 years old
7 (range: 65-91). We collected N= 340 Italian healthy subjects pooled across subgroups for age, sex, and education.
8 Multiple regression analyses were performed to estimate the effect of age, education and sex on the participant's
9 performance. Equivalent scores and cut-off scores were also defined for the number of unique designs (UDs) and the
10 number of strategies (CSs). Results showed that UD increases with decreasing age and increasing educational level.
11 CSs is influenced by higher educational levels but neither by age nor sex. A significant inverse correlation between the
12 UD and percentage of errors occurred, suggesting that a higher number of UD is associated with a fewer number of
13 errors and higher CSs employed. Overall, the MFPT provides a measure of cognitive functioning in terms of the ability
14 to initiate and realize designs, affording useful hints for clinical settings. The MFPT may represent a handy and useful
15 tool with a specific focus in the differentiation of healthy versus pathological aging.

16
17
18
19

20 **Keywords:** Modified Five-Point Test, Figural fluency; design fluency; executive functions; prefrontal cortex;
21 normative data.

22
23
24
25
26

1. INTRODUCTION

Executive functions include different abilities, such as set shifting (i.e., the ability to switch from one mental set to another), control/update representations in working memory, and inhibition of automatic responses [1]. Among the numerous neuropsychological tasks developed to assess executive functions, *fluency tasks* have often been proposed as they examine the ability to maximise the number of responses under constraint of time and restricted search conditions (words or figures) while avoiding response repetition [2]. Fluency tasks have extensively been adopted in neuropsychology under the form of verbal fluency (i.e., the ability to generate multiple words in one minute); specifically, phonemic fluency requires the ability to produce, in one minute, as many unique words as possible starting with a single letter [3][4].

Beside verbal fluency tasks, a non-verbal fluency task has also been developed; here participants are asked to draw as many designs as possible in a time-constrained setting [5]. Non-verbal fluency is referred to the ability to generate unique nonverbal responses by means of cognitive flexibility and is usually related to divergent thinking [6] and creative cognition [7]. Indeed, divergent thinking and creative cognition enhance the ability to quickly initiate and produce a series of possible alternative solutions to tasks that do not require a single correct answer. The higher divergent thinking and creative cognition, the higher are the number of alternative responses. Specifically, divergent thinking involves fluency, flexibility, originality of responses, and productivity measures, meaning that in design fluency, subjects can generate new ideas by inhibiting those previously produced, manipulating information by means of working memory and enhancing the number of unique corrected responses. Spatial cognition, visual attention, and selective attention are also involved in figural fluency tasks' performance. The anatomo-clinical hypothesis beyond the figural fluency tasks is still debated, as, compared to the verbal fluency that results to be impaired due to a left frontal lobe damage, several subsystems in the left and right frontal and posterior regions could lead to deficits in figural fluency tasks [8-10]. The first version of the non-verbal fluency has been the Design Fluency task [5], which includes two parts: (i) a free condition and (ii) a fixed (four-lines) condition. In the free condition subjects have 5 minutes to design different non-sense drawings (i.e., that cannot be named); in the fixed (four-lines) condition, participants have 4 minute' to design many drawings using only four lines. Since the use of this test led to difficulties with scoring (i.e., the comparison of several examiners requires to determine whether one drawing was sense or non-sense, or whether it was the repetition of another one), several versions of free or constrained drawing have been developed to assess figural fluency over the years [11]. An attempt has been made introducing the Five-point test [12]. This test was firstly employed with children between 6 and 13 years old and consists of a sheet on which 40 equal squares are reproduced in a 5 x 8 array, each consisting of five symmetrically arranges dots (Figure 1a). Subjects are asked to connect two or more dots using straight lines by generating as many as possible designs in each rectangle, while, they should avoid repetitions of previous designs. This test had to be completed in 3 minutes of time. The Five-point test, compared to the free designs, has a more structured pattern, is easier to complete and score, and allows to evaluate executive functions with a less involvement of the visuo-constructive and visual-motor skills [13]. The free-drawing condition and the absence of a starting scheme requires a greater effort in activating a mental representation (visuo-constructive and spatial skills) and in translating it into a motor scheme (visuomotor skills) without the aid of visual models within which to organize the drawing. This can lead the subjects to draw very complex objects, involving even more complex visuospatial skills. Moreover, the subject is asked to reproduce unreal drawings, involving a greater contribution of non-verbal, visuo-constructive skills.

1 Afterwards, the Five-point Test proposed by Regard has been then revised by the Ruff's figural fluency test [14]. This
2 test comprises five test conditions of 1 minute each. In each condition, different test material is presented. Although the
3 first condition is analogous to the Five-Points Test, the others either encompass interference patterns or the dots are
4 asymmetrically positioned. The subject is asked to produce different designs utilizing each configuration and has the
5 opportunity to practice prior to each condition. Despite this version is often used, it takes a long time, is redundant and
6 does not provide additional information compared to the form proposed by Regard and colleagues.

7 More recently, Cattelani and co-workers proposed the Modified Five-Point Test (MFPT) [15]. This adaptation is based
8 on Regard's original form, but is simplified, quicker to be administered and easier to understand than previous
9 adaptations. The MFPT takes 3 minutes and includes two pre-test examples that allow participant to get involved in the
10 task and to better understand how to perform it. In fact, the examiner has to exemplify aloud how to connect the dots in
11 each rectangle while drawing himself a possible configuration and while explaining the rules to be complied. This is an
12 advantage compared to the Five-point Test, where the examples are already drawn without any verbal cue.

13 This “modified” version introduced two additional qualitative and quantitative information: the percentage of
14 perseverative and rule-breaking errors considered as an error index (ErrI) and the number of cumulative strategies (CSs)
15 adopted: the number of unique designs incorporated into a strategy (at least three designs organized according to a
16 logical sequence, either enumerative or rotative).

17 According to Ruff (1988), a reduction in the number of production strategies is associated with an increased number of
18 perseverative errors, thus the qualitative analysis of the strategies can help in the differentiation of the patient's
19 performance. However, Cattelani et al., [15] suggested to take also into consideration the percentage of perseverations,
20 as they found that strategy utilization only minimally influenced the production of perseverative errors. Thus, a
21 perseveration index can be suitable to detect a reduced productivity in clinical populations. The normative data
22 published by Cattelani and colleagues have been limited to Italian young and middle age people (range 16-60 years),
23 since their studies were focused on acquired brain-injuries, extensively represented in those age categories (i.e. [16]).
24 However, performance on the figural fluency test has subsequently been shown to be impaired in several
25 neurodegenerative disorders, such as Alzheimer's disease (AD) [17], Parkinson's disease [13], fronto-temporal
26 dementia (FTD) [18], and Multiple Sclerosis [19][20], which involve mainly older people for which the test has not yet
27 been validated. Therefore, the aim of the current study was to collect normative data for older adults (> 60 years old) in
28 a population-based study of aging.

29

30 **2. MATERIAL AND METHODS**

31

32 **2.1 Participants**

33 After obtaining informed consent, all participants involved in a population-based study on aging, named Great-Age,
34 conducted on subjects older than 65 and residents in the city of Castellana Grotte (Bari), in the Puglia Region, in
35 Southern Italy, were enrolled in this study. The detailed methodology of the study has been already published elsewhere
36 [21-23]. In brief, the sample was the electoral list including at December 31, 2014 19,675 subjects, with 4,008 people
37 aged 65 years or older. The neuropsychological assessment for this study began at January 2017 and we cooperated
38 with general practitioners who invited older subjects previously selected to participate in the study with the support of
39 the city census office. A broader description of the recruitment modality it's reported in Figure 2. Participants of the
40 Great-Age study underwent a multidisciplinary assessment, including a neurological and neuropsychological

1 examinations. In particular, the neuropsychological assessment included the MMSE, followed by a comprehensive
2 neuropsychological evaluation. Participants were selected for this study if they fulfilled the following inclusion criteria:
3 (i) no current or past history of major neurological, psychiatric or neurodevelopmental disorders, learning disabilities, or
4 substance abuse; (ii) MMSE corrected score > 24; (iii) no diagnosis of Mild Cognitive Impairment (MCI) according to
5 the DSM – 5 Edition [24]. Figure 2 shows the flow-diagram of the sampling and enrollment process of the GreatAge
6 Study. As per other Italian normative studies (i.e., [25]), to avoid “hyper-normal” sampling group, we did not adopt
7 stringent sampling selection criteria; thus, we did not exclude participants on the basis of presence of highly prevalent
8 chronic medical conditions, like hypertension or type II diabetes, heart failure, or chronic obstructive pulmonary
9 disease.

10 This study was conducted in accordance with the Declaration of Helsinki. The experimental protocol received the
11 ethical approval by the Institutional Review Board of the Institute of Gastroenterology IRCCS “S. De Bellis”. All the
12 participants signed the informed consent prior to their participation in the study.

13

14 **2.2 Modified Five-point test Procedure**

15 The participant is present with sheets of paper on which 40 dot matrices are arranged in an array of eight rows and five
16 columns. In each rectangle (3x2 cm) five dots are drawn (Figure 1c). The examiners asks to produce as many different
17 designs as possible by connecting the dots in each rectangle with one or more straight lines within three minutes. Not all
18 dots per rectangle have to be used. Moreover, the participant is asked to avoid repetition of the same configuration and
19 any designs or draw lines that do not connect dots. Before the beginning of the test, the examiner completes the
20 examples: the first two rectangles at the top of the page, while explaining that it would be possible within each rectangle
21 to draw even one line by connecting only two dots (example of the first rectangle) or more (example of the second
22 rectangle). Furthermore, participants were instructed to continue by filling in each row, from left to right. The number
23 of lines that joined dots were not important, but it was underlined that it would have been essential not to repeat the
24 same design done previously, nor those present in the first two rectangles drawn by the examiner. The test began by
25 setting the 3 minutes of time in which participants could fill in rectangles. When participants completed the first page,
26 the examiner placed the second page next to the first one so that the subject could check it (to avoid perseverations).

27

28 **2.2.1. Scoring**

29 As in Cattelani et al. (2011), four indexes were computed: (i) the number of the designs overall; (ii) the number of
30 errors made including perseverations, lines drawn outside the dots, or not straight lines linking two dots; (iii) the
31 number of unique designs (UDs), obtained subtracting the number of errors from the number of the designs overall; (iv)
32 the number of UD produced within a strategy (CSs). Two possible strategies were considered: enumerative and
33 rotational. The former refers to lines added or subtracted systematically. The latter is counted when the entire figure or a
34 portion of the figure is rotated and the rest of the figure remains fixed. Notably, the use of strategies was considered
35 only when at least three consecutive drawings were completed using the same strategic behavior (i.e., same enumerative
36 or rotative). Further graphical examples of strategic behaviors are reported in Cattelani et al., 2011.

37 As in Cattelani et al., 2011, we considered Error index (ErrI) as the proportion between the cumulative number of failed
38 designs and the number of designs overall.

39

40

1 **2.3 Statistical Analysis**

2 *Demographic Variables:* quantitative variables are reported as mean, SD, median and range, while qualitative variables
3 are reported as percentages. The Spearman correlation coefficient (rho) was used as a measure of association between
4 the quantitative indexes obtained from the MFPT.

5 *Normative data:* to calculate the norms of the MFPT we adopted the equivalent score method proposed by Capitani
6 [26]. Raw scores referring to UDs and CSs were entered into a simple linear regression model. We reduced the residual
7 variance considering the influence of the three main demographic variables (i.e., age, years of education, and sex). The
8 variables found to be significant were then estimated into multiple regression analyses, to generate prediction equations.
9 The final model included only those variables that were deemed significant after multiple regression with the other
10 significant single predictors. All data were considered without any specific treatment for the outlier values, this is
11 mainly due to the fact that the goodness of fit of our model was acceptable. The raw score was then adjusted to remove
12 the influence of age, and of education and sex whether appropriate. Correction grids were generated by reversing the
13 signs of the coefficients to adjust the original score by adding or subtracting the contributions of the significant
14 variables. They were derived only for those scores where the correction factor was > 0.25 , to align the standardization
15 of the MFPT to the previous made by Cattelani et al., 2011.

16 *Equivalent scores (ES)* were computed based on the distribution. First of all, we calculated a cut-off as the outer one-
17 sided non-parametric tolerance limit for the lower 95% of the population with 95% confidence: values equal to/lower
18 (/higher) than the outer tolerance limits were considered as pathological. Then, we estimated 4 cut-off values that
19 divided the original scores into 5 groups: ES = 0, the worst performing group, whose scores were below the 5th
20 percentile of the distribution; ES = 4, the group with normal performance, whose scores were above the median (50th
21 percentile). The remaining 3 groups were derived by segmenting the scores between the 50th percentile and the 5th
22 percentile in 3 segments of equal length. Concerning the CSs variable, a score below the reported cut-off value has been
23 considered as the cut-off of the lowest group and the remaining 3 groups were again derived by segmenting the scores
24 in 3 segments of equal length. A p value $< .05$ was considered as statistically significant. All analyses were carried out
25 using STATA software, version 14 (StataCorp, College Station, Texas).

26

27

28 **3. RESULTS**

29 We enrolled a sample of 340 healthy subjects (218 males). They were native Italian speakers, with a mean age of 71.2
30 (SD: 5.0; range 65–91 years) and a mean of 8.9 years of education (SD: 4.0; range 1-23 years). Twenty-five participants
31 were left-handed. Demographic characteristics of the enrolled participants as long as their distribution according the
32 age, sex and education are reported in Table 1. Demographics variables of the 340 participants and results obtained by
33 the MFPT are detailed in Table 2.

34

35 **3.1 Unique designs (UDs).**

36 Results from the final multiple regression analysis showed that both age and education were significant predictors of
37 UDs performance [$F(3,336) = 26.18$; $p < .001$]. Regression formula for adjusted UDs scores was: $UDs = \text{raw score} -$
38 $[-0.3 \times (\text{age} - 71.2)] + [0.73 \times (\text{education} - 8.9)]$. Age affected negatively the performance, while higher values of
39 education positively influenced performance. Correction grids and ES are shown in Tables 3 and 5. The ES = 0 (i.e., the
40 cut-off score) for UDs was ≤ 10.05 , which identifies an individual's pathological performance.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

3.2 Cumulative strategies (CSs).

Results from the final multiple regression analysis showed that both sex and education were significant predictors of CSs performance [$F(3,336) = 20.19; p < .001$]. Regression formula for adjusted CSs scores was: $CSs = \text{raw score} - [(0.41 \times (\text{education} - 8.9) + 1.16 \text{ (if female)})]$. Education impacted on the number of strategies, with males having a better performance. Correction grids and ES are shown in Tables 4 and 5. The ES = 0 (i.e., the cut-off score) for CSs was 0, i.e., an individual's pathological performance.

3.3 Error index (ErrI).

Failed designs in our sample were perseverative or rule errors. An ErrI score ≥ 0.5 was reported by the worst 5% of the sample subjects, thus the cut-off score indicating that an individual's performance will be abnormal for ErrI is greater than 0.5.

3.4 Correlation analyses.

Correlation analysis between the cumulative number of UD_s and the ErrI scores was found to be statistically significant ($\rho = - 0.217; p < .001$). There was a positive significant correlation between the UD_s and the CS_s ($r = 0.340; p < .001$), and a negative correlation between CS_s and ErrI ($\rho = - 0.251; p < .001$).

4. DISCUSSION

The main result of the present study was to provide norms of the MFPT in a sample of Italian elderly. The MFPT is a modified version of the Five-Point Test. The reference sample of the MFPT consists of Italian subjects aged between 16 and 60 (Cattelani et al., 2011). Reference data of the elderly population can encourage the application of this test in neurodegenerative diseases, which are a frequent condition of subjects aged > 60 years old. Compared with previous studies on figural fluency tests, our results showed that age and education influence the number of unique designs (UD_s), although most of these studies relied on relatively young, highly performing subjects (i.e., [13][15][27-28]). The number of unique designs increases with decreasing age and increasing educational level, whereas sex differences had no influence on UD_s performance. This performance has been described for other figural fluency tests including elderly subjects [i.e., 29]. This might be explained by the fact that participants with higher education can have collected more experience and routine in the execution of executive complex tasks, while aging process can result in a reduced functional connectivity of the dorso-lateral prefrontal cortex [30], that is responsible for the motor planning and to maintain information and goal-directed actions [31]. This reduction in prefrontal cortex might be also responsible for a decrease in executive functions frequently observed in older subjects. However, this is only a hypothesis that needs to be verified by further work. This leads to a limitation of the current study: the lack of MRIs that allow us to make strong statements.

Furthermore, we found that the increase of strategies number (CS_s) was influenced by higher education. This suggests that education influences not only the productivity of designs produced but also the strategies employed. Similarly, Palmiero et al., [32] revealed that lower formal education played a significant role in visual divergent thinking in older people, having an impact on the flexibility component. As stated above in the manuscript, divergent thinking is crucial

1 when older subjects should manipulate new information exploiting working memory abilities. Through flexibility,
2 education can shape the number of unique and alternative responses in a strategic way, as higher educated subjects have
3 experience and routine in execution of executive tasks. Nonetheless, other psychometric measures of executive
4 functions have found a similar trend (i.e., [33; 34]), showing that higher education is linked with better performance in
5 elderly. As strategic behavior hardly ever occurred in lower educated participants, only a simple CSs cut-off score was
6 defined, though, the CSs index can be suitable for clinical application aiming to detect impairment in higher educated
7 patients at early state of the disease (e.g., MCI). In evidence of this, recently, divergent thinking (underling strategic
8 behavior) has been also conceptualized as an indicator of cognitive reserve, a key concept considered a protective factor
9 against cognitive decline, that refers to various factors (e.g., educational level, type of work done, frequency of leisure
10 activities) [35].” Moreover, in contrast with previous studies, we found also a sex influence on the strategy’s
11 performance: in our sample, males are able to produce more strategies than females. These differences can be explained
12 considering the fact that the reference sample of other studies (e.g., [15][36]) included younger people and, in one case
13 [36] did not consider the enumerative, but only the rotative strategies. Compared to the study of Cattelani et al., 2011,
14 our data yield a significant inverse correlation between the UDs and percentage of errors occurred (ErrI), suggesting
15 that a higher number of unique designs is associated to a fewer number of errors, indeed a greater number of UDs is
16 also correlated to a higher number of strategies (CSs) employed, and more strategies (CSs) are associated to a lower
17 number of failed designs (ErrI). As suggested, the number of strategies could influence, at least partially, the
18 performance in terms of the ability to initiate and realize designs [2]; [11][36].

19 The use of MFPT could allow assessing mental productivity, self-monitoring and the ability to regulate responding in
20 the visual–spatial domain. Herby, whether figural fluency is purely a domain of executive functioning or not is still an
21 open question [37]. In terms of underlying brain regions, lesion studies supported that left hemisphere lesions produce
22 more severe deficits in verbal fluency than right-sided lesions, whereas figural fluency is particularly sensitive to right
23 frontal lobe dysfunction [38]. However, Cipollotti et al. (2020) [10] recently showed that figural fluency might have
24 minimal usefulness in the identification of lateralized dysexecutive impairment, as it involves a distributed network not
25 only including frontal regions, but also several subsystems in the left and right hemispheres. On the other hand, verbal
26 fluency seems to be a more appropriate tool to detect localized left frontal regions, thus, they concluded that non-verbal
27 fluency does not mirror the same cognitive and neural processes of verbal fluency.

28 Overall, patients’ performance in figural fluency has shown to be impaired in AD, Parkinson’s disease, MCI, FTD, and
29 multiple sclerosis [18][39-40][21]. In particular, Parkinson’s patients exhibited significant deficits in the number of
30 original designs (productivity) compared to a group of healthy subjects [13]. Moreover, with respect to the other figural
31 fluency tests, the Five-Point Test has shown to be the most suitable measure since it is able to provide different
32 information on perseverative errors and strategic behavior, which seems to be relevant in the differential diagnosis of
33 cognitive impairment in Parkinson’s disease patients compared both to healthy controls and mixed neurological
34 diseases (mostly due to a focal brain damage). Furthermore, FTD patients showed that the number of perseverations on
35 the design fluency task was higher compared to patients with AD [18]. This result was confirmed also by Ranasinghe et
36 al. [41] who revealed that patients with AD are better in monitoring errors than patients with behavioural variant-FTD.
37 However, to date, it is still not sufficiently clear whether figural fluency performance could be relevant and useful as a
38 cognitive marker differentiating among neurodegenerative disease phenotypes. Further psychometric studies are needed
39 to addressed the diagnostic accuracy of such neuropsychological test in neurodegenerative disorders, by assessing also
40 the proportion of patients that can be correctly classified into a specific phenotype, as other Italian normative studies did
41 for other cognitive-linguistic tests (e.g., 42-43)).

1 The normative data presented in this study have three strengths: (i) they were collected on a population-based cohort of
2 a considerable number of elderly subjects over 65 years old (N = 340); (ii) all subjects underwent a comprehensive
3 neuropsychological, neurological and psychiatric assessment, and (iii) our sample was collected in a population-based
4 setting, whereas Italian normative data have usually been based on a convenience sample of healthy individuals. Indeed,
5 the subjects were selected from a population-based study on aging involving a geographically-defined population
6 (residents of a small town of Puglia region, Southern Italy). Furthermore, the test has been administered by only two
7 investigators, reducing the inter-rater variability.

8 A limitation related to our community-dwelling enrollment concerns the unequal distribution of the whole sample
9 stratified by education, age and sex. Since the distribution mirrors the real distribution of the Great-Age population and
10 it is related to the consecutive inclusion of patients, the cells referring to the highest educational level and age (i.e., age
11 80-84 and > 84) result very poor (from 0 to 3 subjects). However, due to the population characteristics, increasing the
12 number of participants with >80 years old will lead to an increase of the number of participants in the 65-79 years old
13 range. We tried to cope with this issue performing a regressive model, which allowed analyzing age and education as
14 continuous variables, avoiding any important missing information in the correction grids. Further studies on the
15 influence of such demographic variables in the oldest old are of course necessary to confirm our results. To wrap up, the
16 population-based design minimizes potential referral bias and enhances the generalizability of the study findings, aimed
17 at recruiting a sample representative of the “real” general population [44][21]. The selection bias might have been
18 avoided in this epidemiological setting as all invited participants were consecutively enrolled, without a-priori
19 excluding any potential healthy subjects. In contrast, convenience sample recruited from the caregivers of patients
20 referring to clinics, hospitals, or subjects from religious communities, recreational centers, educational facilities etc. are
21 by definition willing participants selected with implicit criteria, such as familiarity to patient’s disease, social, economic
22 and educational background providing a narrower spectrum of cognitive functioning, resulting in potentially biased
23 findings coming from different populations. The study is also free from biases due to the use of volunteer samples
24 recruited via advertisement or via referral to specialized centers. Volunteers may participate because of a known family
25 history of the condition under study or may have greater health awareness leading to biased findings.

26 In conclusion, the availability of normative data of the MFPT encourages neuropsychologists to use it in clinical
27 practice. The MFPT could be useful to evaluate dysexecutive deficits, being suitable for diagnostic purposes and in the
28 detection of specific changes, by also bypassing language disorders (i.e., anomias) that could interfere with the
29 detection of the real cognitive abilities of the subject.

30

1 **Tables**

2 **Table 1.** Distribution of the whole sample (N:340) according to stratified by education, age groups and sex. Each
 3 cell reports the number of participants for that specific stratus (F: females; M: males).

Education (years)	Age (years)										Total	4
	65-69		70-74		75-79		80-84		>84			
0-5	M	10	M	17	M	13	M	3	M	1	44	5
	F	24	F	22	F	13	F	2	F	3	64	
6-8	M	42	M	24	M	4	M	3	M	1	74	6
	F	17	F	4	F	3	F	1	F	2	27	
9-13	M	50	M	23	M	8	M	1	M	0	82	7
	F	8	F	7	F	3	F	0	F	0	18	
>13	M	10	M	5	M	3	M	0	M	0	18	8
	F	4	F	4	F	2	F	0	F	3	13	
Total		165		106		49		10		10	340	

10

11 **Table 2.** Demographic features of the enrolled participants and results obtained from the MFPT.

Variables	mean ± SD (%)	Median	Range
Age (years)	71.2 ± 5.0	70	65 to 91
Sex (male)	218 (63.9)	-	
Education (years)	8.9 ± 4.0	8	1 to 23
Total drawings	26.8 ± 10.3	25	7 to 65
UDs Score	21.4 ± 8.2	21	4 to 43
CSs (Count)	2.7 ± 4.9	2	0 to 29
Errors (Count)	5.4 ± 5.4	4	0 to 29
Error index (%)	0.2 ± 0.2	0.15	0 to 0.80

12 Note. Cumulative unique designs (UDs); Cumulative strategies (CSs); Percentage of failed designs (ErrI).

13

14 **Table 3.** Correction grid for cumulative unique designs (UDs)

Education (years)	Age (years)				
	65-69	70-74	75-79	80-84	> 85
5	2.25	3,75	5,25	6,75	8,25
8	0	1,5	3	4,5	6
13	-3,75	-2,25	-0,75	0,75	2,25
18	-7.25	-5.75	-4.25	-2.75	-1.25

15

16

17

1 **Table. 4** Correction grid for cumulative strategies (CSs)

Education (years)	Sex	
	F	M
5	1.5	0.5
8	0.25	-0.75
13	-1.75	-2.75
18	-3.75	-5

2

3 **Table. 5** Equivalent scores for cumulative unique designs (UDs) and cumulative strategies (CSs)

Equivalent score	UDs	CSs
0	≤ 10.05	0
1	10.06-15.30	0.03 - 0.16
2	15.31 - 18.20	0.17 - 0.37
3	18.21- 20.89	0.38 - 0.57
4	≥ 20.90	≥ 0.58

4

5

1 **Figure Caption**

2 **Figure 1** Test material for the three design fluency tests described in the manuscript: a) the Five Point Test sheet
3 (Regard et al., 1982); b) the Ruff's figural fluency test (Ruff et al., 1987); c) the Modified Five-Point Test (the first two
4 dices are used as example; Cattelani et al., 2011).

5
6 **Figure 2.** The flow-diagram of the GreatAge study population showing the enrollment process.
7

8

1 REFERENCES

- 2
- 3 1. Miyake A, Friedman NP, Emerson MJ, et al (2000) The Unity and Diversity of Executive Functions and Their
4 Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cogn Psychol* 41:49–100
- 5 2. Lezak D, Lezak MD, Howieson D, et al (2004) *Neuropsychological Assessment*. Oxford University Press
- 6 3. Milner, B (1964) Some effects of frontal lobectomy in man. *The Frontal Granular Cortex and Behavior* 313–
7 334
- 8 4. Henry JD, Crawford JR (2004) A meta-analytic review of verbal fluency performance following focal cortical
9 lesions. *Neuropsychology* 18:284–295
- 10 5. Jones-Gotman M, Milner B (1977) Design fluency: the invention of nonsense drawings after focal cortical
11 lesions. *Neuropsychologia* 15:653–674
- 12 6. Runco MA (2008) Commentary: Divergent thinking is not synonymous with creativity. *Psychology of*
13 *Aesthetics, Creativity, and the Arts* 2:93–96
- 14 7. Forthmann B, Wilken A, Doebler P, Holling H (2019) Strategy Induction Enhances Creativity in Figural
15 Divergent Thinking. *J Creat Behav* 53:18–29
- 16 8. Robinson G, Shallice T, Bozzali M, Cipolotti L (2012) The differing roles of the frontal cortex in fluency tests.
17 *Brain* 135:2202–2214
- 18 9. Schmidt CSM, Nitschke K, Bormann T, et al (2019) Dissociating frontal and temporal correlates of
19 phonological and semantic fluency in a large sample of left hemisphere stroke patients. *Neuroimage Clin* 23:101840
- 20 10. Cipolotti L, Molenberghs P, Dominguez J, et al (2020) Fluency and rule breaking behaviour in the frontal
21 cortex. *Neuropsychologia* 137:107308
- 22 11. Ruff RM, Light RH, Evans RW (1987) The ruff figural fluency test: A normative study with adults. *Dev*
23 *Neuropsychol* 3:37–51
- 24 12. Regard M, Strauss E, Knapp P (1982) Children’s production on verbal and non-verbal fluency tasks. *Percept*
25 *Mot Skills* 55:839–844
- 26 13. Goebel S, Atanassov L, Köhnken G, et al (2013) Understanding quantitative and qualitative figural fluency in
27 patients with Parkinson’s disease. *Neurol Sci* 34:1383–1390
- 28 14. Evans RW, Ruff RM, Gualtieri CT (1985) Verbal fluency and figural fluency in bright children. *Percept Mot*
29 *Skills* 61:699–709
- 30 15. Cattelani R, Dal Sasso F, Corsini D, Posteraro L (2011) The Modified Five-Point Test: normative data for a
31 sample of Italian healthy adults aged 16–60. *Neurol Sci* 32:595–601
- 32 16. Tibæk M, Forchhammer HB, Dehlendorff C, et al (2017) Incidence and mortality of acquired brain injury in
33 young Danish adults between 1994 and 2013: a nationwide study. *Brain Inj* 31:1455–1462
- 34 17. Fama R, Sullivan EV, Shear PK, et al (2000) Structural brain correlates of verbal and nonverbal fluency
35 measures in Alzheimer’s disease. *Neuropsychology* 14:29–40
- 36 18. Possin KL, Chester SK, Laluz V, et al (2012) The frontal-anatomic specificity of design fluency repetitions and
37 their diagnostic relevance for behavioral variant frontotemporal dementia. *J Int Neuropsychol Soc* 18:834–844
- 38 19. Hansen S, Muenssinger J, Kronhofmann S, et al (2017) Cognitive screening in Multiple Sclerosis: the Five-
39 Point Test as a substitute for the PASAT in measuring executive function. *Clin Neuropsychol* 31:179–192
- 40 20. Pitteri M, Ziccardi S, Dapor C, et al (2019) Lost in Classification: Lower Cognitive Functioning in Apparently
41 Cognitive Normal Newly Diagnosed RRMS Patients. *Brain Sciences* 9:321

- 1 21. Tortelli R, Lozupone M, Guerra V, et al (2017) Midlife Metabolic Profile and the Risk of Late-Life Cognitive
2 Decline. *Journal of Alzheimer's Disease* 59:121–130
- 3 22. Lozupone M, Panza F, Piccininni M, et al (2018) Social Dysfunction in Older Age and Relationships with
4 Cognition, Depression, and Apathy: The GreatAGE Study. *J Alzheimers Dis* 65:989–1000
- 5 23. Sardone R, Lampignano L, Guerra V, et al (2020) Relationship between Inflammatory Food Consumption and
6 Age-Related Hearing Loss in a Prospective Observational Cohort: Results from the Salus in Apulia Study. *Nutrients*
7 12.: <https://doi.org/10.3390/nu12020426>
- 8 24. American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*.
9 American Psychiatric Pub
- 10 25. Siciliano M, Chiorri C, Passaniti C, et al (2019) Comparison of alternate and original forms of the Montreal
11 Cognitive Assessment (MoCA): an Italian normative study. *Neurol Sci* 40:691–702
- 12 26. Capitani E, Laiacona M (1997) Composite neuropsychological batteries and demographic correction:
13 Standardization based on equivalent scores, with a review of Published Data. *J Clin Exp Neuropsychol* 19:795–809
- 14 27. Izaks GJ, Joosten H, Koerts J, et al (2011) Reference data for the Ruff Figural Fluency Test stratified by age
15 and educational level. *PLoS One* 6:e17045
- 16 28. Tucha L, Aschenbrenner S, Koerts J, Lange KW (2012) The Five-Point Test: Reliability, Validity and
17 Normative Data for Children and Adults. *PLoS ONE* 7:e46080
- 18 29. Ross TP, Foard EL, Hiott FB, Vincent A (2003) The reliability of production strategy scores for the Ruff
19 Figural Fluency Test. *Archives of Clinical Neuropsychology* 18:879–891
- 20 30. Zhu Z, Johnson NF, Kim C, Gold BT (2015) Reduced frontal cortex efficiency is associated with lower white
21 matter integrity in aging. *Cereb Cortex* 25:138–146
- 22 31. Szczepanski SM, Knight RT (2014) Insights into human behavior from lesions to the prefrontal cortex. *Neuron*
23 83:1002–1018
- 24 32. Palmiero M, Nori R, & Piccardi L. (2017) Verbal and visual divergent thinking in aging. *Experimental brain*
25 *research*, 235(4), 1021-1029.
- 26 33. Zalonis I, Kararizou E, Triantafyllou NI, Kapaki E, Papageorgiou S, Sgouropoulos P, et al. (2007) A normative
27 study of the Trail Making Test A and B in Greek adults. *The Clinical Neuropsychologist*, 22, 842–850.
- 28 34. Van der Elst W, van Boxtel MPJ, van Breukelen, GJP, & Jolles J. (2006). The Stroop Color-Word Test:
29 Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, 13,
30 62–79.
- 31 35. Stern Y, (2009) Cognitive reserve. *Neuropsychologia*, 47(10), 2015-2028.
- 32 36. Goebel S, Fischer R, Ferstl R, Mehdorn HM (2009) Normative data and psychometric properties for qualitative
33 and quantitative scoring criteria of the Five-point Test. *Clin Neuropsychol* 23:675–690
- 34 37. Ghanavati E, Salehinejad MA, Nejati V, Nitsche MA (2019) Differential role of prefrontal, temporal and
35 parietal cortices in verbal and figural fluency: Implications for the supramodal contribution of executive functions. *Sci*
36 *Rep* 9:3700
- 37 38. Foster PS, Williamson JB, Harrison DW (2005) The Ruff Figural Fluency Test: heightened right frontal lobe
38 delta activity as a function of performance. *Arch Clin Neuropsychol* 20:427–434
- 39 39. Peter J, Kaiser J, Landerer V, et al (2016) Category and design fluency in mild cognitive impairment:
40 Performance, strategy use, and neural correlates. *Neuropsychologia* 93:21–29

- 1 40. Kramer JH, Quitania L, Dean D, et al (2007) Magnetic resonance imaging correlates of set shifting. *J Int*
2 *Neuropsychol Soc* 13:386–392
- 3 41. Ranasinghe KG, Rankin KP, Lobach IV, et al (2016) Cognition and neuropsychiatry in behavioral variant
4 frontotemporal dementia by disease stage. *Neurology* 86:600–610
- 5 42. Battista P, Catricalà E, Piccininni M, et al (2018) Screening for Aphasia in NeuroDegeneration for the
6 Diagnosis of Patients with Primary Progressive Aphasia: Clinical Validity and Psychometric Properties. *Dement Geriatr*
7 *Cogn Disord* 46:243–252
- 8 43. Pigliautile M, Chiesi F, Primi C, et al (2019) Validation study of the Italian version of Communication
9 Activities of the Daily Living (CADL 2) as an ecologic cognitive assessment measure in older subjects. *Neurological*
10 *Sciences* 40:2081–2088
- 11 44. Roberts RO, Geda YE, Knopman DS, et al (2008) The Mayo Clinic Study of Aging: Design and Sampling,
12 Participation, Baseline Measures and Sample Characteristics. *Neuroepidemiology* 30:58–69