



# UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale

Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata - FISPPA

Corso di Laurea Magistrale in Clinical, Social, and Intercultural  
Psychology

Tesi di Laurea Magistrale

Valutazione Cognitiva Non Verbale (NoVA): proprietà psicometriche  
e confronto tra parlanti nativi e non nativi di lingua italiana

Non-Verbal Cognitive Assessment (NoVA): psychometric properties and comparison  
between native and non-native Italian speakers.

*Relatrice*

Prof.ssa Sara Mondini

*Correlatore*

Dott. Andrea D'Ippolito

*Laureando: Gabriele Sissa*

*Matricola: 2102422*

Anno Accademico 2025/2026



INDEX	
ABSTRACT	4
CHAPTER 1	
1.1: Language as a Methodological Variable in Neuropsychological Assessment	5
1.2: Non-Verbal and Language-Reduced Cognitive Assessment: Potentials and Limitations	7
1.3: Cognitive Assessment in Native and Non-Native Speakers	10
1.4: Study rationale and objectives	14
CHAPTER 2	
2.1. NoVA: Non-Verbal Assessment Battery Test	15
2.2 Tele-Global Examination of Mental State (Tele-GEMS)	35
CHAPTER 3	
3.1: Data collection for the Normative Sample	38
3.2 Administration of NoVA Battery to Non-Native speakers	58
DISCUSSION AND CONCLUSIONS	68
REFERENCES	74
APPENDIX	82

## **Abstract**

The present thesis is part of the development of a new non-verbal instrument for cognitive assessment, the Non-Verbal Assessment Battery (NoVA). It focuses on the analysis of the preliminary normative data collected from a sample of healthy Italian participants.

NoVA is a screening battery for global cognitive functioning composed of 16 subtests exploring different domains, including attention, visuospatial exploration, praxis, visuo-constructive abilities, short- and long-term visual memory, and executive functioning. The data collection was carried out on 162 healthy Italian native speakers, aged between 18 and 88 years, with an education range between 5 and 27 years.

Cognitive Reserve was also considered through the short version of the Cognitive Reserve Index questionnaire (sCRIq, Mondini et al., 2023), with a score value ranging from 68 to 169.

A multiple linear regression analysis was conducted to examine the effect of sociodemographic variables, such as age, education, and cognitive reserve, on the global NoVA score. The results showed that these variables significantly influenced the performance, explaining about 60% of the variance of the total score with a  $p < 0.001$ . Furthermore, the battery also showed a satisfactory internal consistency (raw Cronbach's  $\alpha = 0.59$ ; st  $\alpha = 0.81$ ).

In order to provide a preliminary support for the construct validity of the battery, the relationship between NoVA and another established instrument for neuropsychological screening – such as Tele-GEMS (Montemurro et al., 2023) – was also examined, showing a moderate positive correlation ( $r = 0.528$ ).

Finally, an exploratory analysis was carried out on a small group of 20 non-native speakers, in order to obtain indications of the possible impact of the linguistic barrier on the administration of the battery.

Overall, the results suggest that the NoVA Battery represents a promising tool for cognitive assessment in contexts where language acts as a confounding factor. At the same time, further studies on larger and more heterogeneous samples will be necessary to better define the battery's psychometric properties.

## **CHAPTER 1**

### **1.1: Language as a Methodological Variable in Neuropsychological Assessment**

Neuropsychological assessment is a fundamental clinical and methodological tool to investigate the relationship between brain functioning and behaviour (Harvey, 2012). Broadly, neuropsychological evaluation aims to identify strengths and patterns among specific cognitive domains (e.g., memory, attention, executive functioning, language, perception, and reasoning) (Lezak, 2004). The assessment is typically conducted through structured test batteries, whose purpose is to evaluate cognitive performance, support diagnostic formulation, and rehabilitation from an objective perspective. Neuropsychological testing is particularly relevant when changes are observed in cognitive and behavioural functioning, such as alterations in memory, reasoning, concentration, or personality, that might reflect a neurological or systemic medical condition (Bennett, 2001).

Clinicians can use standardised instruments to compare individual performance with normative data; this process facilitates the identification of discrepancies from the expected cognitive functioning, enabling the monitoring of cognitive change over time (Lezak, 2004; Harvey, 2012). Nevertheless, despite the use of these standardised procedures, neuropsychological assessment cannot be considered completely independent from linguistic influences.

Even when tests do not specifically target language skills, language still plays a role in various testing stages. These stages include explaining and understanding task instructions, processing verbal or verbally delivered stimuli, and generating a response, particularly when time is limited (McLaughlin & Nayak, 1989).

The language in which a test is administered has been shown to influence performance, even in neurologically intact bilingual speakers, suggesting that linguistic tasks affect cognitive outcomes independently of actual cognitive abilities (Gasquoine et al., 2007; Lehman Blake et al., 2015; Kassir et al., 2025). Additionally, research on measurement equivalence across languages has highlighted how differences in linguistic and cultural contexts might influence the psychometric characteristics of neurocognitive instruments, increasing concerns about the comparability of the scores obtained across different language groups (Goodman et al., 2021). Recent studies

have emphasised the importance of rigorous statistical validation when adapting cognitive tests across languages and cultures, ensuring that the same instruments measure the same construct across populations (Daga et al., 2025). Moreover, emerging research suggests that examiner-related linguistic factors, like subtle variations in verbal interaction while testing, may play a role in influencing cognitive task performance (Li et al., 2025).

Taken together, these findings indicate that language should not be considered an irrelevant variable in neuropsychological assessment, but a methodological component that systematically influences test performance and test interpretation. This issue becomes even more important when assessing people who are tested in a non-native language: here, the performance may reflect not only cognitive functioning itself but also linguistic load and familiarity with the testing context.

From both theoretical and clinical perspectives, culture and language are now recognised as major variables influencing neuropsychological test performance. Cognitive assessment tools that have been developed within specific linguistic and cultural contexts might show reduced accuracy when administered to participants from different backgrounds, particularly in specific conditions, like early cognitive impairment, where performance differences are often subtle but clinically present (Ng et al., 2018). Beyond the linguistic component, factors such as cultural familiarity, educational experiences, and exposure to assessment settings may influence performance independently of actual cognitive functioning.

From a clinical standpoint, cognitive assessment cannot be completely neutral: performance is always mediated by the interaction between the examiner, the examinee, and the testing context (Hodges, 2017). Numerous factors, such as instructions' comprehension, stimulus interpretation, and expectations concerning test-taking behaviour, may affect both the performance and the interpretation of the results.

Therefore, clinicians are required to interpret neuropsychological scores considering the clinical and socio-cultural context of the individual, rather than relying solely on performance and its relative comparison with normative data.

In response to these challenges, increasing research efforts have focused on the development of language-neutral cognitive assessment tools. For example, the Visual

Cognitive Assessment Test (VCAT; Kandiah et al., 2016) has been specifically designed to minimise linguistic and cultural influences on the examinee. It demonstrated good validity across multiple Southeast Asia populations, supporting the feasibility of cross-cultural cognitive screening approaches (Lim et al., 2018).

These findings underline the importance of considering both linguistic and cultural variables in cognitive assessment, supporting the growing interest in assessment approaches that reduce language-related bias.

## **1.2: Non-Verbal and Language-Reduced Cognitive Assessment: Potentials and Limitations**

Non-verbal cognitive assessment refers to the evaluation of cognitive functioning through tasks designed to minimise the influence of linguistic processing on the participant's performance (Botting, 2005). Non-verbal assessment approaches have been developed because researchers recognised that many traditional cognitive tests rely heavily on language comprehension and verbal production, potentially limiting their applicability in populations with language impairments or that are culturally and linguistically diverse. Early efforts to reduce such linguistic bias led to the development of language-free intelligence measures. For example, the Test of Nonverbal Intelligence (TONI; Brown, 2003) was specifically designed to assess reasoning abilities without requiring verbal responses and with minimal language comprehension (Brown, Sherbenou, & Johnsen, 1990; Brown, 2003).

From a broader clinical perspective, cognitive evaluation is inherently multidimensional and involves integrating verbal, nonverbal, and behavioural information (Reinhard et al., 2008; Shaughnessy & Weintraub, 2025). Clinical assessment models underline that cognitive functioning cannot be fully assessed through verbal performance alone, as non-verbal communication, behavioural responses, and execution strategies are also fundamental to provide relevant diagnostic information (Nardone et al., 1992; Gkintoni et al., 2024). In this regard, non-verbal cognitive assessment represents an important attempt to isolate cognitive processes from linguistic influences, allowing an accurate estimation of cognitive functioning in situations where language represents a potential confounding factor.

Within a theoretical framework, non-verbal assessment has been increasingly conceptualized as a tool to improve fairness and accessibility in cognitive testing, especially in multicultural and multilingual contexts (De Meuse et al., 1987; Mancuso et al., 2024). Recent frameworks on non-verbal assessment have emphasized how important it is to reduce language-based tasks to obtain more accurate representations of cognitive functioning in individuals with either limited language proficiency or communication difficulties (McCallum, 2003; Wasserman, 2017). This perspective aligns with the rising recognition that neuropsychological assessment needs a balance between measurement validity and cultural/linguistic fairness.

From a clinical perspective, the use of non-verbal (or low-language) measures has shown particular relevance in conditions where language impairment might interfere with the accurate overall evaluation of cognitive domains. For example, research that compared verbal and non-verbal memory measures in neurodegenerative conditions has shown that language deficits may confound the interpretation of verbal test performance, while non-verbal tasks may provide complementary information useful for differential diagnosis (Butler et al., 2009).

Building on the need to reduce linguistic bias in neuropsychological evaluation, several attempts have been made to develop cognitive tools with minimal verbal demands. Nevertheless, despite the development of non-verbal or language-reduced cognitive tools, the total elimination of linguistic and cultural influences remains an important challenge (Rosselli & Ardila, 2003; Cormier et al., 2022). Familiarity with testing procedures, educational background, and culturally shaped problem-solving strategies may still indirectly affect test performance, even when the tasks are designed to minimize verbal demands (Ardila, 2007). Differences in educational quality, exposure to testing environments, and opportunities for visuospatial learning may influence performance in non-verbal cognitive measures, rather than exclusively reflecting underlying functioning (Rosselli & Ardila, 2003). In addition, cognitive evaluation is inherently multidimensional, involving not only cognitive processing but also examiner-examinee interaction, communication dynamics, and contextual interpretation of task demands (Nardone et al., 1992).

Through a clinical lens, non-verbal cognitive tests have shown utility in populations with language impairments, such as aphasic patients, where traditional language-

mediated assessment may underestimate preserved reasoning and problem-solving abilities (Christy & Friedman, 2005). Instruments specifically designed to reduce linguistic load, such as the Test of Nonverbal Intelligence (Brown et al., 2003), represent an important effort in differentiating and dissociating cognitive abilities from language proficiency. However, recent literature emphasizes the importance of rigorous cross-cultural and demographic validation procedures: large-scale neuropsychological investigations have demonstrated that cognitive assessment must be validated across populations to preserve diagnostic accuracy and fairness (Franzen et al., 2022).

Furthermore, the increasing use of digital cognitive assessments has introduced additional variables, like digital familiarity, accessibility, and user-interface comprehension, which can influence performance in certain populations, for example, in older adults (Cubillos & Rienzo, 2023).

Collectively, these findings suggest that while non-verbal and language-reduced cognitive assessments represent an important step toward a more objective neuropsychological evaluation, they need to be interpreted within a broader biopsychosocial and cultural framework, especially when cognitive tools are designed for use across heterogeneous linguistic, cultural, and educational backgrounds. This perspective is relevant when selecting assessment tools in clinical contexts involving linguistically heterogeneous populations, highlighting the importance of choosing instruments that balance ecological validity and diagnostic sensitivity.

Under these limitations, modern neuropsychological research emphasizes the importance of developing assessment tools that focus equally on management of linguistic load preservation and ecological and construct validity. Instead of aiming for complete language neutrality (which can be theoretically unrealistic), current approaches focus on minimizing unnecessary verbal mediation, maintaining sensitivity to the cognitive processes measured. This perspective aligns with a broader movement focused on culturally informed neuropsychological assessment, where demographic variability, educational exposure, and sociocultural background are considered during test development as integral components of cognitive evaluation, rather than confounding variables to eliminate (Ng et al., 2018; Lim et al., 2018). In line with this perspective, further investigations showed that many

neuropsychological instruments are strongly rooted in Western educational models and literacy assumptions. This phenomenon may limit their diagnostic accuracy and comparability across populations with different backgrounds (Franzen et al., 2022).

Within this view, visual and cognitive tasks have gained growing relevance, especially in multicultural and multilingual contexts. In this regard, non-verbal cognitive assessment represents an important attempt to reduce the linguistic demands associated with traditional neuropsychological testing, allowing a more accurate estimation of cognitive functioning in situations where language might act as a confounding factor (Brown et al., 1990; Lezak, 2004). However, language processing remains tightly interconnected with attentional and executive systems that support comprehension, rule maintenance, working memory, and response selection during testing (Baddeley, 2003; Hodges, 2017). For this reason, the influence of language cannot be completely eliminated from neuropsychological evaluation (Rosselli & Ardila, 2003; Ardila, 2005).

However, as recent cross-cultural studies highlight, even visually based tools require careful normative calibration and population-specific validation procedures to ensure diagnostic accuracy and fairness in different groups (Goodman et al., 2021; Daga et al., 2025).

From a clinical perspective, this evolution marks a shift from viewing language as a variable to be removed to understanding it as a component of a broader cognitive and environmental system. As a consequence, the selection of neuropsychological instruments needs to consider not only psychometric validity, but also contextual applicability, particularly in settings where traditional test standardization might not reflect the person's characteristics, or when assessments are conducted in linguistically heterogeneous populations (Chaytor & Schmitter-Edgecombe, 2003).

### **1.3: Cognitive Assessment in Native and Non-Native Speakers**

#### **1.3.1: Linguistic Diversity in Clinical Contexts**

The issue of language in cognitive assessment has been particularly relevant because of the progressive increase in linguistic diversity within clinical and research settings. Taking the Italian context as an example, neuropsychological evaluations are

frequently conducted with participant who, despite living in Italy, might not have Italian as their first language or might present differentiated levels of proficiency. In these situations, cognitive testing may be influenced not only by neurological functioning but also by linguistic abilities, familiarity with testing environments, and educational background. The diagnostic interpretation might be affected by test comprehension and response strategies, which are influenced by cultural and linguistic contexts embedded in cognitive assessment tools (Ng et al., 2018). Research has also shown that these variables may influence performance, independently of the examinee's underlying cognitive functioning (Rosselli & Ardila, 2003). These challenges have been increasingly recognised at the European level, where the population is becoming increasingly heterogeneous, highlighting the need for careful interpretation of cognitive test results and for assessment strategies to consider background diversity (Nielsen, 2022).

This introduces an important challenge: low performance in cognitive testing might reflect true cognitive impairment but might also derive from increased cognitive load associated with processing and thinking in a non-native language. Investigations have demonstrated that scores on neuropsychological testing are strongly associated with neurological status, as well as with ethnicity, linguistic abilities, and acculturation variables, suggesting that failing to consider these factors may lead to methodological misinterpretation (Boone et al., 2007; Razani et al., 2007). The need for careful interpretation of results has recently been demonstrated in linguistically heterogeneous European populations: non-native speakers show lower performance on neuropsychological measures, even in the absence of neurological disease (Stålhammar et al., 2022). For these reasons, it is important to understand how native versus non-native language speaker status is associated with neuropsychological performance. This is a necessary step toward improving the accuracy and fairness of instruments for cognitive assessment in increasingly multicultural populations.

### 1.3.2: Cognitive Processing in a Non-Native Language

Growing literature suggests that differences in performance between native and non-native speakers may be due to how the linguistic information is processed during task execution. People operating in a non-native language may adopt different processing

strategies, for example, increased attention to linguistic input and greater reliance on contextual interpretation, which may influence the performance. Communication research has highlighted the use of compensatory processing strategies in cognitively demanding situations; listeners often adjust their processing when interacting with non-native speakers, allocating attentional resources to interpret meanings under conditions of uncertainty (Long, 1996; Lev-Ari, 2015). Performing tasks in a non-native language might require the simultaneous management of linguistic processing and task-solving processes. This dual goal can increase reliance on the executive control system, especially concerning rapid rule maintenance, working memory, and response monitoring (Hodges, 2017). These effects may also emerge in neurologically healthy participants, especially in tasks requiring comprehension and in time-constrained tasks (Ng et al., 2018).

### 1.3.3: Diagnostic Validity and Risk of Misinterpretation in Non-Native Speakers

An emerging body of literature suggests that neuropsychological test performance is not always representative of underlying cognitive functioning, especially when participants take the test in a non-native language. Performance on neuropsychological measures is only partially explained by general intellectual functioning: test scores reflect interacting cognitive and contextual processes, rather than pure cognitive ability (Diaz-Asper et al., 2004). Cross-linguistic research has consistently shown that language plays a significant role in performance between native and non-native speakers (Carone et al., 2006). Linguistically mediated tasks tend to show major differences between native and non-native speakers (even if they are conversationally fluent), while tasks involving the visuo-constructive skills are the ones that show minor differences (Stålhammar et al., 2022). These performance discrepancies carry important diagnostic consequences. False positives and misdiagnoses may occur when neuropsychological tests are administered to people from diverse linguistic and cultural backgrounds (Norman et al., 2000; Norman et al., 2011). The risk of diagnostic misclassification increases when neuropsychological instruments developed in specific cultural and linguistic contexts are applied to culturally diverse individuals, especially when normative samples do not represent the assessed population (Daugherty et al., 2017). Clinical literature similarly emphasizes

the role of language, history, educational background, and sociocultural context in neuropsychological assessment as a fundamental requirement to prevent underestimation of impairment (Harris et al., 2003).

Beyond linguistic variables, broader determinants such as socioeconomic and psychosocial factors (Yaffe et al., 2013; Sisco et al., 2015; Zahodne et al., 2017), have been shown to significantly contribute to variability in cognitive performance across populations (Zahodne, Sol, et al., 2019). Disparities in cognitive test results are partially explained by contextual life variables, including education quality, stress exposure, and social opportunity structures (Zahodne et al., 2021). Clinical research in Europe also highlighted that culturally diverse aging populations require assessment tools that explicitly account for the examinee's educational levels, as the language in which cognitive tests are administered, as well as the level of formality used, shows a significant impact on the communication with the examiner, and its subsequent test score (Ardila, 2005). and address the migration-related experiences to preserve diagnostic validity (Nielsen, 2022).

Emerging research emphasizes how important ecological validity is in real-life contexts as well as the role of language experience when interpreting cognitive performance. Studies have examined how heritage and multilingual populations demonstrate that tests based on proficiency measures often fail to capture the complexity of real-world language and cognitive functioning (Gullifer et al., 2020; Luque et al., 2025). Cross-cultural investigations, including research on migrant assessed on a neuropsychological level outside their country of origin, underline how culturally shaped cognitive strategies and life experiences can influence performance patterns (Fratti et al., 2011; Ardila, 2021). These findings highlight the need for cognitive assessments to differentiate between language-mediated performances and their underlying cognitive functioning, improving the development of ecologically valid and culturally based assessment approaches. The development of assessment frameworks that underline reduced linguistic confounding while preserving clinically sensitive variations represents a critical priority in today's research. Such approaches aim to improve diagnostic precision and to support equitable access to neuropsychological evaluation across subjects from diverse populations.

#### **1.4: Study rationale and objectives**

The considerations discussed throughout this chapter highlight the importance of developing neuropsychological instruments that minimize linguistic mediation, while preserving sensitivity to the cognitive processes being evaluated. In increasingly multilingual and culturally heterogeneous contexts, cognitive test performance might be influenced not only by neurological functioning but also by educational background, sociocultural experience, and language proficiency. Therefore, instruments developed to reduce linguistic demands represent an important approach for improving the fairness and interpretability of neuropsychological evaluations.

Within this framework, the Non-Verbal Assessment (NoVA) battery was developed as a cognitive assessment tool aimed at evaluating multiple cognitive domains while limiting the potential confounding effect of language. The battery relies primarily on visually mediated tasks and reduced verbal requirements to provide a more accessible evaluative tool for cognitive functioning across individuals with different linguistic and cultural backgrounds.

Specifically, the present study aims to:

- examine the preliminary psychometric properties of the NoVA Battery in a normative Italian sample;
- through regression analyses, evaluating the influence of sociodemographic variables, namely age, education, and cognitive reserve, on global performance
- provide preliminary evidence of construct validity by examining the relationship between NoVA performance and an established neuropsychological screening tool (Tele-GEMS);
- explore the feasibility of administering the NoVA Battery to a non-native population, in order to investigate the potential impact of reduced linguistic demands on test performance;
- compare native and non-native participants, through the creation of homogeneous subgroups matched for age, education, and cognitive reserve, in order to better isolate the role of language in cognitive assessment.

## CHAPTER 2

### 2.1. NoVA: Non-Verbal Assessment Battery Test

Following is a list of tests that compose the updated protocol for the NoVA Battery (Appendix):

- Visual-Spatial Attention:
  - o Flowers Barrage
  - o TMT-A (Trail-Making Test A)
  - o TMT-B (Trail-Making Test B)
  - o Symbol Figure Modalities Test
- Visuospatial Construction
  - o Copy of a bi-dimensional drawing
  - o Copy of a three-dimensional drawing
- Motor programming
  - o Hand movements imitation
  - o Motor sequence imitation
- Visual Agnosia
  - o Incomplete figures recognition
- Memory
  - o Immediate visual memory
  - o Delayed visual memory
  - o Visual Colour Span – Forward
  - o Visual Colour Span – Backwards
- Reasoning
  - o Logical relations
  - o Interference
  - o Go/No-Go

Some subtests require stimulus presentation via PC, using a timed PowerPoint presentation that automatically displays the involved stimuli (e.g., Interference, Go/No-Go, Visual Colour Span, etc.). Administration of the entire battery takes approximately 40 minutes. For each subtest, a raw score is obtained and subsequently weighted so that the score of each subtest contributes equally to the overall battery score. In the following paragraph, all subtests currently included in the battery will be

presented. For clarity, the subtests will not be described following the administration order but will instead be organized according to the cognitive domain assessed. Each subtest description will also include the instructions provided to the examinee and the criteria used for score attribution.

## **Visual-Spatial Attention**

Within the NoVA Battery, visual-spatial attention abilities are assessed through the administration of four subtests.

### **Flowers Barrage**

The flower's barrage subtest is designed to assess visual-spatial search abilities and psychomotor speed. This is an adaptation of Gauthier et al., *Bells Test* (1989), in which targets and distractor stimuli have been replaced with different figures. In this task, the examinee is required to cancel the target symbols by drawing a line through them, among a series of distractors presented on the paper sheet. Although the spatial distribution of the target symbols may appear random, they are equally distributed in the two halves of the sheet.

#### **Administration**

The examiner presents the sheet to the examinee and provides the following instructions: "*On this paper, there are various symbols. Your task is to cross out all the flowers you see, as shown in the example above*" (the examiner points at the image on top of the sheet). "*Be careful, because you will have one minute, so try to cross out as many flowers as possible before the time runs out. Let's begin*". The examiner starts the stopwatch at the beginning of the task.

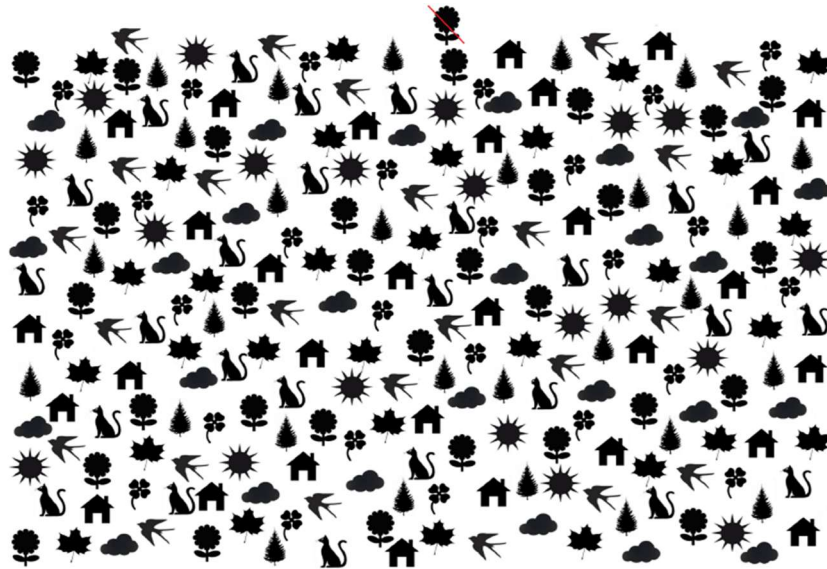


image 2.1: Flowers barrage test sheet

## SCORING.

The score corresponds to the number of target symbols correctly cancelled within 60 seconds. The maximum obtainable score is 30.

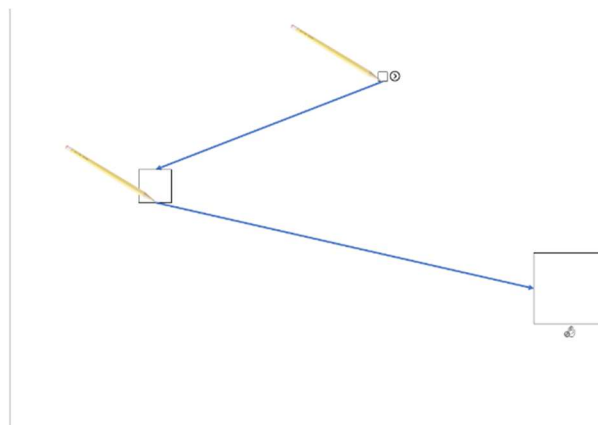
## Trail Making Test (TMT-A)

The Trail Making Test – A (TMT-A) subtest is designed to assess visual-spatial search abilities, selective attention, and psychomotor speed. It represents an adaptation of Reitan’s (1958) original version, in which the numerical stimuli have been replaced with geometric shapes to reduce linguistic and numerical demands. In this task, the examinee is required to connect a series of squares increasing in size by drawing a continuous line, starting from the smallest square and proceeding sequentially to the largest one. To facilitate the task’s comprehension and provide an example of correct execution, a PowerPoint presentation is used for the Trail Making subtests (both TMT-A and TMT-B). The presentation displays a simplified version of the task and illustrates the correct connection procedure.

### Administration.

The examiner first presents the demonstrative PowerPoint presentation to the examinee and provides the following instruction: “*On the screen, you can see squares*

*of different sizes. They are connected from the smallest to the largest with a line.”* Subsequently, the examiner presents a practice sheet and gives the following instructions: *“On this sheet, you can see squares of different dimensions as well. Your task is to connect the squares from the smallest one (pointing at the starting square) to the largest (pointing at the final square) by drawing a line. Do not need to worry if the line is not perfectly straight. Try to complete the task without lifting the pencil from the sheet.”*



*Image 2.2: TMT-A demonstrative slide from PowerPoint*

After ensuring that the examinee understood the task, the examiner presents the test sheet and states: *“This is the actual test. As before, you will need to connect the squares from the smallest to the largest, from the beginning to the end, by drawing a line. This time, I will record the time, so try to be as fast as you can. You may start whenever you are ready.”*

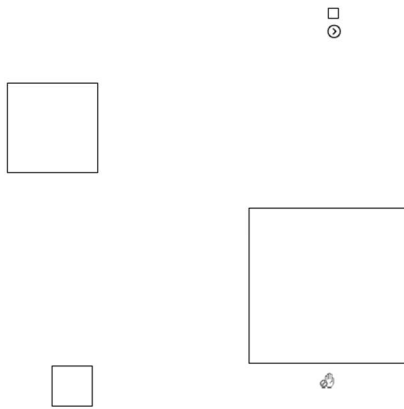


Image 2.3 TMT-A practice sheet

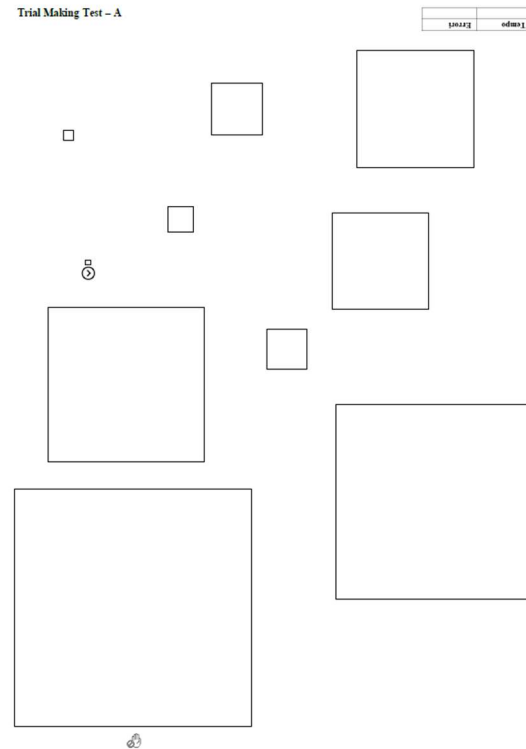


Image 2.4 TMT-A test sheet

### Scoring.

The score corresponds to the number of correct connections made by the examinee. If the examinee commits an error, the examiner does not interrupt the task and allows the examinee to continue. Subsequent correct connections are counted as valid. The maximum obtainable score is 9.

### Trail Making Test (TMT-B)

The Trail Making Test – B (TMT-B) subtest is designed to assess alternating and selective attention, visual-spatial search abilities, and psychomotor speed. It represents an adapted version of Reitan's test (1958), and of the *Colour Trails Test* (CTT) included in the *CNTB* battery test (Nielsen et al., 2018). Compared to TMT-A, this task is more complex, as it requires a double alternation rule: the examinee is required to connect, by drawing a line, squares and coloured circles (blue and green) presented on the paper sheet, following a square-circle-square-circle sequence. The squares increase progressively in dimension, while the circles alternate in colour (blue-

green-blue-green).

To facilitate the task's comprehension, a PowerPoint is used to demonstrate the alternation rule and the correct execution of the task.

Administration.

The examiner first presents the demonstrative PowerPoint presentation and provides the following explanation: *"Here you can see squares (pointing) of different sizes and coloured circles: some are blue (pointing), and some others are green (pointing). As you can see, (examiner starts the animation), the squares alternate with the circles (pointing) and gradually increase in size. The circles alternate in colour, following the order blue-green-blue-green. Therefore, you will connect a square (pointing) to a blue circle (pointing), then a larger square (pointing), then to a green circle (pointing), and so on until the end (pointing)"*.

The examiner then presents the practice sheet and states: *"On this sheet, you can see squares (pointing) of different sizes, as well as blue and green circles (pointing). By drawing a line, you must alternate squares and circles, starting from the smallest square (pointing), increasing to the largest one. The circles, on the other hand, must alternate in colour, blue-green-blue-green. For example, from the smallest square (pointing), you will reach the blue circle (point and track), from there you will connect to a larger square (point and track), and then to the green circle (point and track).*

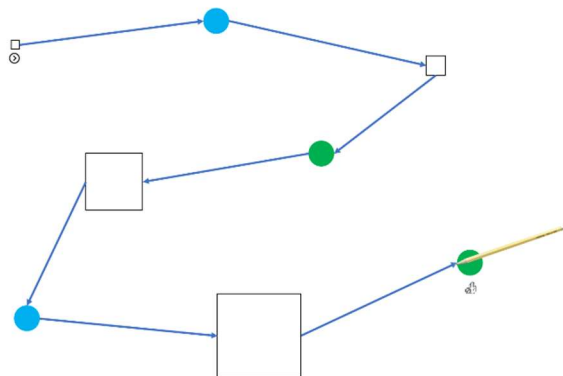


Image 2.5: TMT-B demonstrative slide from PowerPoint

After ensuring that the task has been understood, the examiner presents the test sheet and provides the following instructions: *"This is the actual test, and as you can see, there are more squares and circles than in the practice sheet. As before, you will*

alternate squares and circles by drawing a line, starting from the smallest square (pointing). Remember the squares increase in size, and the circles alternate in colour, starting with green. This time I will track the time, so try to be as fast as possible. You may begin when you are ready.”

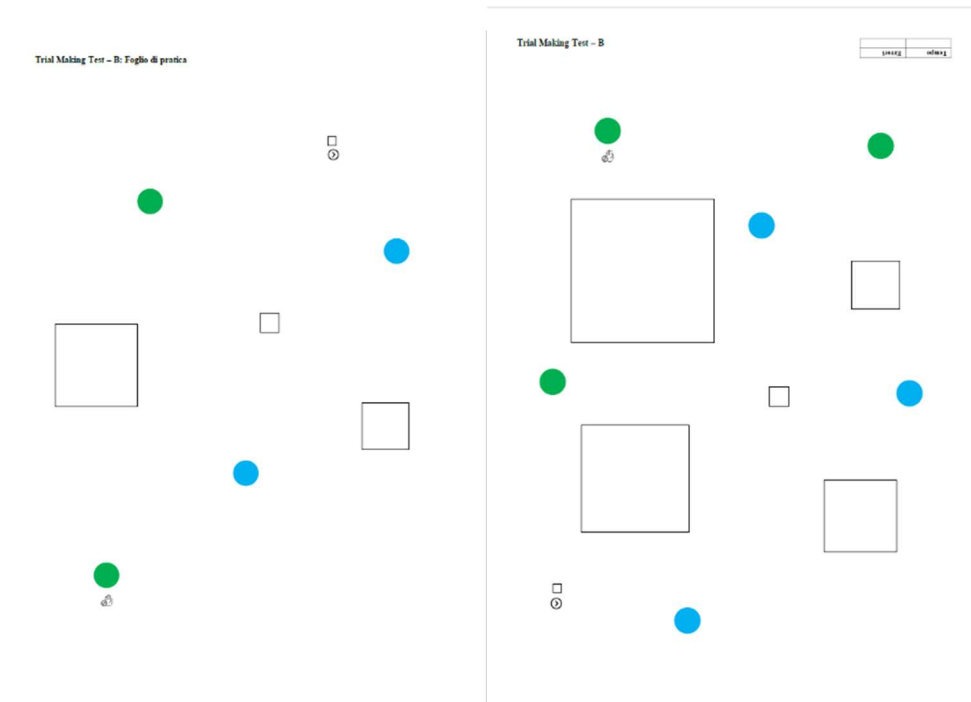


Image 2.6 TMT-B practice sheet

Image 2.7 TMT-B test sheet

### Scoring.

The score corresponds to the number of correct connections made by the examinee. In the event of an error, the examiner does not interrupt the task and allows the examinee to continue. Subsequent correct connections are counted as valid. The maximum obtainable score is 11.

### Symbol Figure Modalities Test (SFMT)

The Symbol Figure Modalities Test (SFMT) is designed to assess psychomotor speed and visual-spatial attention. It represents an adapted version of the *Symbol Digit Modalities Test* originally developed by Smith (1973). Unlike the original version, this adaptation is based on associations between symbols and figures, as indicated in the legend displayed at the top of the sheet.

Administration.

The examiner presents the test sheet to the examinee and provides the following instructions: “Here you can see several figures (pointing at the silhouette), each associated with a specific symbol (pointing at the symbols). For example, the first figure (pointing) is associated with this symbol (pointing), the second figure (pointing) is associated with this other symbol (pointing), and so on. Your task is to fill in the empty spaces (pointing) with the correct symbol corresponding to each figure. For example, for the first figure (pointing at the bird), you will draw the associated symbol (pointing at the X). For the second figure (pointing at the house), you will draw this other symbol (pointing at the O). Now it’s your turn. When the examinee fills the first line, the examiner will say: “Well, from now on, you will have 90 seconds to fill in the paper sheet. Try to make as many associations as possible.”

**Symbol Figure Modalities Test**  
Qui si sono varie figure (silhouette) e con esse associati dei simboli (indici). Come vede alla prima figura (indica) è associato questo simbolo (indica), alla seconda (indica) questo simbolo (indica) e così via. Quello che deve fare è scrivere negli spazi vuoti (indica) il simbolo corrispondente alla figura. Per esempio, per la prima figura (indica) dovremo scrivere questo simbolo (indica X) e ancora, per la casa (indica) dovremo scrivere questo simbolo (indica O e ancora). Adesso provi lei. Avanzate pochi centimetri a disposizione 90 secondi, quindi deve cercare di essere il più veloce possibile. Iniziamo.


---


Image 2.8: SFMT test sheet

NB: the examinee needs to proceed from left to right, completing the task line by line without skipping any spaces.

## Scoring

The score corresponds to the total number of correct symbol-figure associations completed within 90 seconds. The maximum obtainable score is 54.

## Visual-Spatial Construction

Within the NoVA Battery, visual-spatial construction abilities are assessed through two drawing tasks of increasing difficulty: the examinee reproduces two geometric figures. The stimuli to be copied are presented on the administration protocol sheet.

### Copy of a Bi-Dimensional Drawing

The copy of a Bi-dimensional drawing subtest is designed to assess visual-perceptive and visual-constructive abilities. It represents an adapted version of the drawing task included in the *Mini-Mental State Examination* (MMSE; Folstein et al., 1983), in which the examinee is required to draw the intersection between two pentagons; while in the NoVA battery, the task has been replaced with two hexagons.

#### Administration

The examiner presents the stimulus figure to the examinee and states: “*Can you see this figure (pointing)? Try to draw it in the space provided here (pointing to the blank space on the side) as accurately as you can.*”

#### Scoring

The maximum obtainable score is 2 points, assigned if both the hexagons are drawn with six sides each, and if the intersection is correctly represented. A score of 1 point is given if both hexagons have six sides, but the intersection is missing, or if the intersection is correctly drawn but the hexagons do not have six sides. A score of 0 points is given if the hexagons do not have six sides, and the intersection is not correctly drawn.

## Copy of a Three-Dimensional Drawing

The copy of a three-dimensional drawing subtest is designed to assess visual-perceptive and visual-constructive abilities. The subtest is an adapted version of the *Three-Dimensional Drawing Task* included in the *Esame Neuropsicologico Breve – 3* (ENB-3; Mondini et al., 2022), in which the examinee is required to draw a house. In the NoVA Battery, the original task has been replaced with a three-dimensional figure resembling a horse's hoof.

Administration.

The examiner presents the stimulus to the examinee and states: “*Can you see this figure?* (pointing). *Try to copy it in the space provided here* (pointing to the blank space on the side) *as accurately as you can.*”

Scoring

The maximum obtainable score is 2 points, assigned if the drawing looks exactly like the stimulus. A score of 1 point is given if the examinee recognizes the model but omits or inaccurately reproduces some elements (up to two sides, the perspective, or mirror orientation). A score of 0 points is given if the examinee is unable to recognize the model.

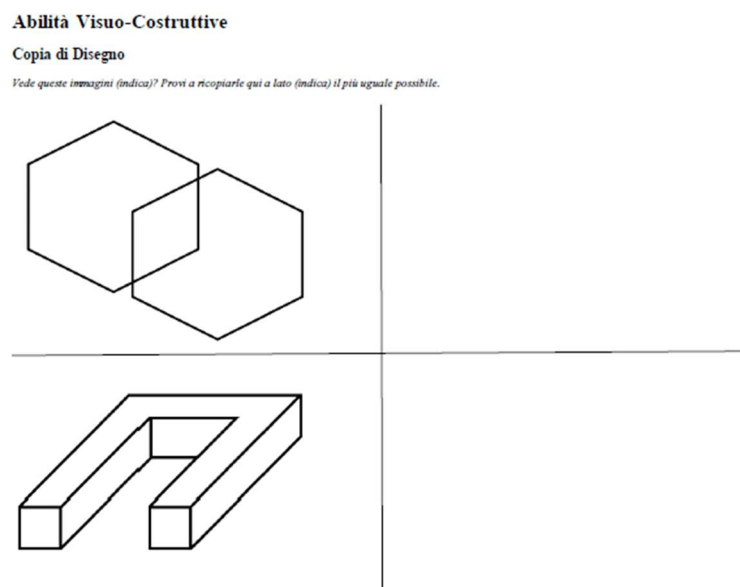


Image 2.9 Bi-dimensional and three-dimensional drawing spaces

## Hand Movements and Motor Sequences

Within the NoVA Battery, motor skills abilities are assessed through two subtests that require the examinee to reproduce hand movements and motor sequences shown by the examiner.

### Imitation of a Hand Movement With No Meanings

The Imitation of Meaningless Hand Movements subtest is designed to assess ideomotor praxis, requiring the examinee to imitate hand movements demonstrated by the examiner. The task is adapted from the *NEPSY-II battery* (Korkman et al., 2007).

#### ADMINISTRATION.

The examiner demonstrates a meaningless hand movement and instructs the examinee as follows: *“I will show you some movements with my hand that you will need to imitate. For example, try to replicate what I am doing (show the first hand movement)”*.

The examiner ensures that the examinee understood the task, and then continues: *“Well, let’s continue in the same way. If I do this (shows the second item), what should you do?”* The procedure is repeated for all remaining items.

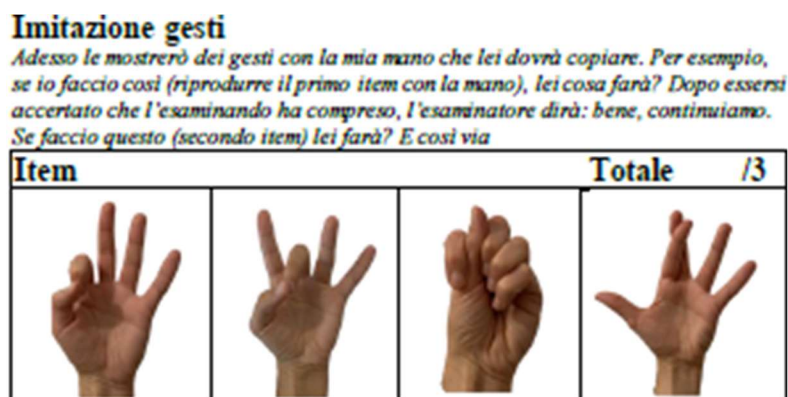


Image 2.10 Hand movement imitation items

#### SCORING.

One point is assigned for each movement correctly reproduced. The maximum obtainable score is 3.

*NB:* The first item is used for demonstration purposes and is not included in the scoring.

## **Motor Sequence Imitation**

The Motor Sequence Imitation subtest assesses the ability to reproduce structured motor sequences demonstrated by the examiner. The sequence corresponds to the original motor program proposed by Lurija, which consists of three consecutive movements: fist-HORIZONTAL; cut-VERTICAL; palm-HORIZONTAL WITH PALM ORIENTED DOWNWARD. The test consists of imitating this three-movement sequence.

Administration.

The examiner provides the following instructions: *“I will show you a sequence of movements that you need to carefully observe. After you have observed them, we will perform the movements together. Then you will continue to perform it on your own until I tell you to stop. Please, start performing the movements only when I tell you to do so.* The examiner demonstrates the sequence three times during the observation phase (the examiner sits still). Afterward, the examiner and the examinee perform the sequence together three times. Finally, the examinee is asked to continue performing the sequence independently six times (until the examiner says *“Stop”*).

Scoring.

The scoring is based on the criteria proposed by Dubios et al. (2000). A score of 3 points is assigned if the examinee performs the sequence correctly and independently for six repetitions. A score of 2 points is assigned if the examinee performs at least three consecutive series independently. A score of 1 point is assigned if the examinee makes errors during the independent execution, but no errors during the combined execution. A score of 0 points is assigned if the examinee is unable to perform three consecutive series together with the examiner.

## **Visual Agnosia**

Within the NoVA Battery, visual agnosia is assessed through a subtest aimed at evaluating integrative visual processing abilities.

## Incomplete Figures Recognition

The Incomplete Figures Recognition subtest is designed to assess the integrative visual agnosia. It represents an adapted version of *InComplete Letter* task from the *Visual Object and Space Perception Battery* (VOSP; Warrington & James, 1991). In this task, an incomplete target image is presented, and the examinee is required to identify the corresponding complete figure, among the given alternatives.

### Administration

The examiner presents a demonstrative slide to the examinee and provides the following instructions: “*On the screen, you can see an incomplete image (points). Your task is to choose which of the images shown below (points) corresponds to the complete version of the one on top. For example, does this image (points at the first image on the left) look like this one (points at the incomplete image)? Does this image (points at the second image on the left) look like this one (points at the incomplete image)? Does this image (points at the first image on the right) look like this one (points at the incomplete image)? Does this image (points at the second image on the right) look like this one (points at the incomplete image)?* The examiner ensures that the examinee understands the task before proceeding.

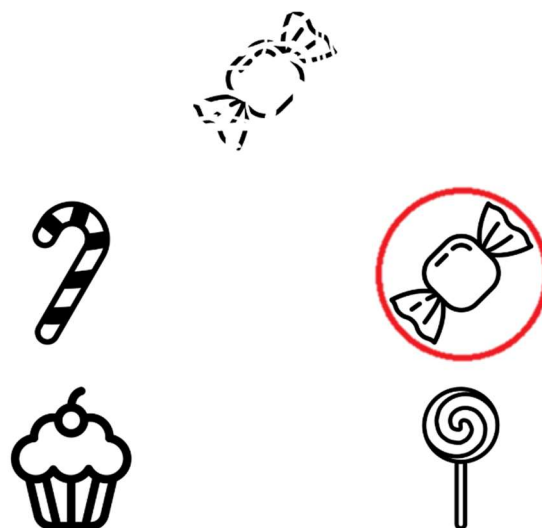


Image 2.11 incomplete figures demonstrative slide

### Scoring

One point is assigned for each correct answer. The maximum obtainable score is 4 points.

## Memory

Within the NoVA Battery, mnemonic functioning is assessed through four tasks. Two tasks investigate visual recognition memory, both immediate and delayed, while two additional tasks assess immediate visual memory and working memory.

### Visual Memory Test - Immediate

The following subtest is designed to assess immediate visual recognition memory. The task is initially presented on a PC's monitor; subsequently, a set of four images is displayed (3 distractor images and 1 target image). The examinee is required to identify the image corresponding to the previously presented target. Between the presentation of the target image and the response slide, four alternatives, an interference screen is displayed to reduce the persistence of the mnemonic trace. The subtest was developed taking inspiration from the face recognition task of the *Rivermead Behavioural Memory Test – Third Edition* (RBMT-3; Wilson, 2008), while its procedural structure is inspired by the *Free and Cue Selective Reminding Test* (FCSRT; Buschke, 1984).

#### Administration

The examiner presents the slide containing the first target image (square with a circle and a triangle within) and provides the following instructions: "You need to memorise the image I am about to show you. After that, I will show you some more figures, and you need to point at the one you just memorised. Can you see this image (the first target image is shown)? Well, memorise it (the examiner clicks on the button to start the animation). Tell me which of these is the one you just saw. Now we will continue.

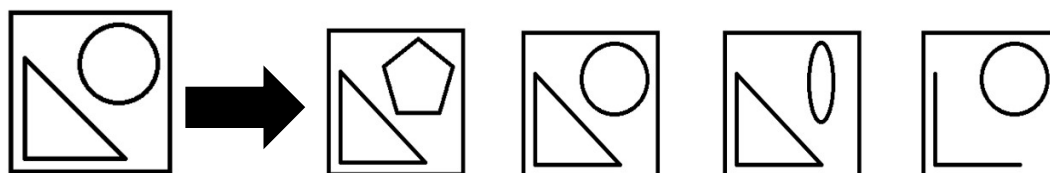


Image 2.12 First item of the immediate memory and immediate recognition screen

#### Scoring

One point is assigned for every correctly recognised image. The maximum obtainable score is 10. The examiner needs to note the number of errors committed.

## **Visual Memory test – Delayed Recognition**

The Visual Memory test – Delayed Recognition subtest is designed to assess long-term visual recognition memory. The task represents an adapted version of the face recognition test included in the *Rivermead Behavioural Memory Test – Third Edition* (RBMT-3; Wilson, 2008). In this task, the examinee is required to recognise, among a set of distractor images, the target images previously shown during the Delayed Visual Memory Test - Immediate. Right after the completion of the Immediate Recognition task, all target images are presented again, one at the time, to be encoded for subsequent delayed recognition. The Delayed Recognition Test is administered after a retention interval of approximately 10 minutes. During this time, tasks that do not involve perceptually similar stimuli are administered to minimize interference effects.

### Administration

Immediately after the Visual Memory Test – Immediate, the examiner presents the target images and provides the following instructions: *“I will now show you some images that you need to remember. Try to fix them in your mind, because later I will ask you to recognise them among a series of images. Let’s start.”* Each target image is displayed for approximately 3 seconds.

After a 10-minute retention interval, during which non-interfering tasks are administered (hand-gesture imitation, motor sequence, reasoning, incomplete figures, interference, go/no-go, colour digit forward and backward), the examiner provides the following instructions: *“Think about the images I asked you earlier to memorise. Now I will show you a sequence of images, and for each one, you will need to tell me if it was among the ones you were asked to remember. Let’s start* (the examiner shows the images and for each one asks the examinee if the image is among the target ones).

### Scoring

One point is given for each correctly recognised image. The maximum obtainable score is 10. The examiner needs to note the number of errors committed.

## Visual Colour Span - Forward

The Visual Colour Span – Forward subtest is designed to assess short-term visual memory. The task represents an adapted version of *Wechsler's Memory Scale* (1945). Traditionally, in digit-span tasks, the examinee is required to remember and immediately repeat a sequence of numbers, in the same order in which they are presented, with the sequences progressively increasing in length. In the NoVA Battery, in order to reduce the influence of the verbal and numerical components, digits have been replaced with coloured squares, which the examinee does not need to name, but simply to indicate by pointing. A PowerPoint presentation is used to present the stimuli, which displays one square per second, ensuring standardized timing and the correct colour order. The presentation includes buttons with specific functions to support the administration (one to press if the examiner recognises the sequence, and another one in case of errors).

### Administration

The examiner presents a demonstrative slide and provides the following instructions: *“Now I will show a sequence of colours. After you have seen the colour sequence, I will show an image with several colours. You will have to point at the colours in the same order as they were shown. We will start with a short sequence, and then the number of colours will increase. Let's give it a try. If I show you these colours (examiner starts the animation), which colours will you point to (click on the icon)? Good, now we can continue.* After the examiner shows two practice examples, the actual task begins. The examiner proceeds with a longer sequence length only if the examinee correctly reproduces the sequence, or if the examinee makes an error, but correctly reproduces the second sequence of the same length.

### Scoring

The score corresponds to the maximum length of the last colour sequence correctly reproduced by the examinee. The task is interrupted after two consecutive errors. The maximum obtainable score is 8 (corresponding to the longest colour sequence presented).

## Visual Colour Span - Backward

The Visual Colour Span - Backward subtest is designed to assess short-term memory and working memory abilities. As the previous task, this subtest represents an adapted version of *Wechsler's Memory Scale* (1945). The task shares the same structure and stimulus configuration as the Visual Colour Span – Forward, the key difference is that after the stimulus presentation, the examinee is required to reproduce the colour sequence in reverse order, starting from the last colour present to the first. As in the previous task, for a correct administration, stimulus presentation is standardized through a PowerPoint presentation, with the same characteristics (visual and timing parameters) used for the Visual Colour Span – Forward.

### Administration

The examiner presents a demonstrative slide and provides the following instructions: *“Now I will show a sequence of colours. After you have seen the sequence, I will show an image with several colours. You will need to point at the colour in reverse order than they were shown, starting from the last colour and going back to the first one. Let's try with an example, and keep in mind that you need to start from the last colour you see and name them in the backward order. If I show you these (starts the animation), which colour will you point to (click on the icon)? Good, now we can continue.*

After the examiner shows two practice examples, the actual task begins. The examiner proceeds to a longer sequence only if the examinee correctly reproduces the sequence, or if the examinee makes an error on the first attempt but correctly reproduces the second sequence of the same length.

### Scoring

The score corresponds to the maximum length of the last colour sequence correctly reproduced by the examinee. The task is interrupted after two consecutive errors. The maximum obtainable score is 8 (corresponding to the longest colour sequence presented).



Image 2.13 colour sequence screen

## Reasoning

Within the NoVA Battery, the cognitive domain of Reasoning is assessed through the administration of three subtests.

## Logical Relations

The Logical Relations subtest is designed to assess abstraction and reasoning abilities. This task represents an adapted version of the Matrix Reasoning task from the *Wechsler Adult Intelligence Scale – Fourth Edition* (WAIS-IV; Wechsler, 2008). For each item, the examiner is required to identify the logical association rule among the elements presented in a grid and to select the stimulus that correctly completes the empty spot, from a set of alternatives presented below the grid.

### Administration

The examiner displays a demonstrative slide and provides the following instructions: *“On the screen, you can see a grid divided into four spaces. Three of the spaces are filled with images, while one is empty. Your task is to identify which of the images below (points) correctly completes the grid. Pay attention, because you need to understand the correct association rule that links the images already present. For example, here we have a rabbit (points), a carrot (points), and a monkey (points). Which image should go in the empty spot (points)? The banana (examiner starts the animation). We can tell it is the correct one because rabbits (points) eat carrots (points), so monkeys (points) eat bananas. Let’s continue.”*

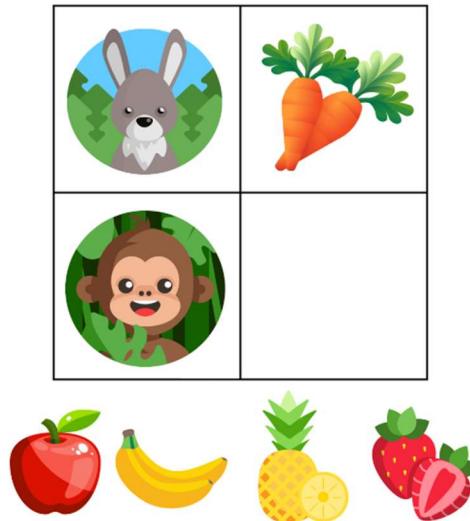


Image 2.14 Logic relations demonstrative slide

### Scoring

One point is assigned for each correct answer. The maximum obtainable score is 4.

### Interference

The Interference subtest is designed to assess attentional control and sensitivity to interference. The task represents an adapted version of the *Frontal Assessment Battery* (FAB; Dubois et al., 2000) task that evaluates interference sensitivity. In the original version, the examiner taps once or twice on the table, the examinee is required to answer back inversely (tapping once when the examiner taps twice and tapping twice when the examiner taps once). In the NoVA Battery, this norm is adapted into a non-verbal visual format: the examinee is required to tap once when a black circle is displayed, and to tap twice when a red circle is displayed.

### Administration

The examiner presents a demonstrative slide and provides the following instructions: *“If you see a black circle (points), tap your hand on the table once (examiner demonstrates and asks the examinee to do it as well). If you see a red circle (click the right button on the mouse and points), tap your hand on the table twice (the examiner demonstrates and asks the examinee to do it as well). Good, now let’s put everything together: if you see a black circle, tap once; if you see a red circle, tap twice. Let’s*

*start.*"

The examiner starts the animation with the left mouse button.

#### Scoring

A total of 16 stimuli are presented. Scoring is based on the number of correct responses as follows:

- 4 points: no errors (16/16 correct)
- 3 points: 1-4 errors
- 2 points: 5-8 errors
- 1 point: 9-12 errors
- 0 points: 13 or more errors

The maximum obtainable score is 4.

### **Inhibition – Go/No-Go**

The Inhibition – Go/No-Go subtest assesses attentional capacity and inhibitory control. The task is also an adaptation of the *Frontal Assessment Battery* (FAB; Dubois et al., 2000) Go/No-Go. In this version, the examinee is required to tap the hand on the table twice when a black circle is displayed on the screen, and to inhibit the motor response when a red circle appears (not knocking at all). The examiner explicitly updated the instructions from the previous Interference task to minimize carry-over effects.

#### Administration

The examiner presents a demonstrative slide and provides the following instructions: *"If you see a black circle (points), tap your hand on the table twice (the examiner demonstrates and asks the examinee to do it as well). If you see a red circle (click on the left mouse button and point), do not tap. Good, now we put the two things together: tap twice when you see a black circle and do not tap when you see a red circle. Let's start."*

#### Scoring

A total of 16 stimuli are presented. Scoring is based on the number of correct responses as follows:

- 4 points: no errors (16/16 correct)

- 3 points: 1-4 errors
- 2 points: 5-8 errors
- 1 point: 9-12 errors
- 0 points: 13 or more errors

The maximum obtainable score is 4.



*Image 2.15 black and red circles from Inhibition and Go/No-Go tests*

## **2.2 Tele-Global Examination of Mental State (Tele-GEMS)**

The Tele-Global Examination of Mental State (Tele-GEMS) is a neuropsychological screening instrument developed to assess global cognitive functioning through a standardized and structured procedure (Montemurro et al., 2023). The instrument was designed to provide an extensive evaluation of multiple cognitive domains. A key feature of the instrument is its remote administration format, which makes it suitable for both clinical and research settings.

The instrument covers multiple cognitive domains, including temporal and spatial orientation, immediate and delayed memory, working memory, attention, language, executive functioning, and abstract reasoning. The subtests target abilities such as naming, verbal comprehension, verbal fluency, metaphor comprehension, spatial representation, and auditory attention. Overall, it offers a broad overview of the examinee's cognitive functioning.

Two parallel versions of the instrument are available: Version A and Version B. The two versions are equivalent in structure and cognitive areas assessed but differ in stimulus content. In the present study, each participant was administered only one version of the instrument.

#### Administration

Administration follows the standardised procedures described by Montemurro et al. (2023). Prior to tele-GEMS administration, the short version of the Cognitive Reserve Index Questionnaire (sCRIq; Nucci et al., 2012) was administered. The assessment must be conducted in a quiet environment, minimizing external distractions. The examiner ensured that the examinee did not have any hearing impairment that could invalidate the test, that the internet/telephone connection was stable, and that the examinee was in a suitable environment, minimizing external distractions. Participants were asked to remove writing materials and to conceal clocks or calendars if present in the room. The order of the subtests is maintained according to the standardized protocol.

#### Scoring

Each Tele-GEMS subtest yields a partial score; a global cognitive score is obtained by summing the individual subtest scores. Total scores were calculated using a standardized Excel file provided by the authors through the Open Science Framework (OSF), in accordance with the official scoring guidelines.

Tele-GEMS was included to provide a standardized measure of global cognitive functioning, allowing comparison with performance on the NoVA Battery and supporting the interpretation of domain-specific results within a broader cognitive framework.


<b>Tele-G</b> lobal <b>E</b> xamination of <b>M</b> ental <b>S</b> tate <b>A</b>		 University of Padua HIT - Human Inspired Technology Research Centre Montemurro et al., (2023)
DATE: _____ Time: _____ Examiner: _____ Name and Surname: _____ AGE: _____ EDUCATION: _____ Handedness: <input type="checkbox"/> L <input type="checkbox"/> R <input type="checkbox"/> AMB CRI: _____ CRI-Educator: _____ CRI-Work: _____ CRI-LeisureTime: _____		
<b>1) ORIENTATION</b> <span style="float: right;">1 point for each correct answer - max 4</span> points Temporal orientation: What season are we in? _____ [ ] Spatial orientation: What year is it? _____ [ ] Relative to Europe, is Africa located to the North, South, East or West? _____ [ ] NOTES _____ TOTAL [.../4]		
<b>2) IMMEDIATE RECALL</b> <span style="float: right;">1 point for each correct word - max 6 points</span> Now I will say six words (one a second); please, listen carefully to all of them and then repeat them aloud (the order does not matter). Memorise these words because I will ask you to repeat them again in a few minutes. The six words are: (1) SUN [ ] (2) NAIL [ ] (3) MOUTH [ ] (4) ARCH [ ] (5) LIGHTHOUSE [ ] (6) MEADOW [ ] NOTES _____ TOTAL [.../6]		
<b>3) MONTHS BACKWARDS / WORKING MEMORY</b> <span style="float: right;">1 point for each correct answer - max 3 points</span> Now, please list the months of the year backwards, starting from December and going back two months at a time, which means you should always skip one month. Therefore, December, October... and so on until I say 'stop'. AUG [ ] JUN [ ] APR [ ] FEB [ ] DEC [ ] NOTES _____ TOTAL [.../5]		
<b>4) SPATIAL REPRESENTATION</b> <span style="float: right;">1 point for each correct answer - max 4 points</span> Now I will ask you to imagine a clock. I will say some specific times and for each of them you should say if the two hands of the clock are positioned: both on the right side, both on the left side, or one in each half of the clock face. (1) At 3:45 the minute hand and the hour hand are positioned: right [ ] left [ ] both [ ] (2) At 9:50 the minute hand and the hour hand are positioned: right [ ] left [ ] both [ ] (3) At 4:20 the minute hand and the hour hand are positioned: right [ ] left [ ] both [ ] (4) At 1:15 the minute hand and the hour hand are positioned: right [ ] left [ ] both [ ] NOTES _____ TOTAL [.../4]		
<b>5) NAMING</b> <span style="float: right;">1 point for each correct answer - max 4 points</span> Now I will read some descriptions and you should tell me what they refer to. For example: "Something you put on your finger when you get married. What is it?" (wedding band - wedding ring - ring) Well, let's start: (1) Animal with black and white stripes that lives in the Savannah _____ (zebra) (2) Building with towers where kings and queens live _____ (castle) (3) Tool used to draw perfect circles _____ (compass) (4) Part of the face that allows you to see _____ (eye/eyes) NOTE _____ TOTAL [.../4]		
<b>6) DELAYED RECALL</b> <span style="float: right;">1 point for each correct word - max 6 points</span> A short while ago I read six words, which you then repeated. Do you remember them? (1) SUN [ ] (2) NAIL [ ] (3) MOUTH [ ] (4) ARCH [ ] (5) LIGHTHOUSE [ ] (6) MEADOW [ ] NOTES _____ TOTAL [.../6]		

Image 2.16 tele-GEMS test sheets

### 7) COMPREHENSION

Please listen carefully to the following instructions and do what is requested of you: Say the letter A twice after saying the letter B once. Suggestions and repetitions of the instructions are NOT allowed. 1 point if the execution is correct - max 1 point  
 [ ] Letter B [ ] Twice letter A

NOTES \_\_\_\_\_ TOTAL [.../3]

### 8) AUDITORY ATTENTION

Now I will read a series of numbers and every time you hear '2' you must say the word 'GREEN', and every time you hear '4' you must say the word 'RED'; when I read other numbers, you must not say anything. So, if I say '2', you say ... and if I say '4', you say ... Let's try: 2, 3, 6, 2, 4. If understood, go on with the test reading a number every two seconds. 1 point for each correct answer - Max 3 points

2	6	7	4	9	1	3	2	2	9	8	5	6
3	2	4	6	7	9	1	4	7	6	8	9	4

NOTES \_\_\_\_\_ TOTAL [.../8]

### 9) VERBAL FLUENCY

Now, I would like you to say in one minute as many words as you can think of (avoiding personal names and city names) which begin with the letter 'T'. An example is 'table'. (If they say proper nouns, remind them to avoid these categories, without stopping the timer). 1 point for each correct word (in max 1 minute

1	5	9	13	17	21	25
2	6	10	14	18	22	26
3	7	11	15	19	23	27
4	8	12	16	20	24	28

NOTES \_\_\_\_\_ TOTAL [.....]

### 11) METAPHOR COMPREHENSION

Now I'll read you a sentence and three possible explanations for it. Please, choose the most appropriate. 1 point if the answer is correct

- Today I visited the city library. That archive is a mine!
- That archive contains precious jewels
- That archive contains interesting documents
- That archive is in the city centre

NOTES \_\_\_\_\_ TOTAL [.../1]

Please use the Excel File present in the Open Science Framework (OSF) to calculate the global score.

## **CHAPTER 3**

This chapter describes the procedures adopted for data collection and the analysis conducted in order to investigate the psychometric properties of the NoVA Battery.

The study was primarily based on a normative sample of healthy Italian participants to collect preliminary normative data and examine the key characteristics of the instrument. Additionally, an exploratory analysis was conducted on a smaller group of Non-Native speakers in order to investigate the applicability of the battery in contexts characterised by linguistic barriers.

The present chapter presents the characteristics of the normative sample, the instruments used, and the procedures adopted for data collection. Subsequently, the analysis conducted on the normative dataset is described, followed by the results obtained from the administration of the battery to the Non-Native sample. Finally, an analysis based on two homogeneous subgroups is presented in order to allow a more controlled comparison between the Native and Non-Native participants.

### **3.1: Data collection for the Normative Sample**

In this section, the characteristics of the normative sample and the procedures used for data collection are presented. The sample includes 162 healthy Italian participants. This phase of the study aimed to collect preliminary normative data for the NoVA Battery (Appendix) in order to examine the psychometric properties of the instruments.

When developing a new neuropsychological assessment tool, the collection of normative data represents a crucial step, as it allows clinicians and researchers to interpret the performances that will emerge when the instrument is used in clinical practice by comparing future examinees with the scores obtained from a reference population.

Indeed, by collecting data from a sample that is representative of the target population (in this case, the Italian population) and by taking into account relevant demographic variables, such as age, education, and cognitive reserve, it is possible to obtain score distributions that support the interpretation of test performance of future examinees. The following sections describe the procedures adopted for data collection and present the results obtained from the normative sample. Subsequently, the results

merging from the administration of the battery to a smaller sample composed of non-native Italian speakers will also be discussed, in order to explore the feasibility of administering the test battery to people with different linguistic backgrounds and potential barriers. Data collection for both samples was conducted between October 2025 and February 2026.

### 3.1.1 Sample

The sample involved in the study consists of  $N = 162$  healthy Italian people, including 91 women and 71 men, with an age range between 18 and 88 years ( $M = 47.123$ ;  $SD = \pm 18.425$ ). The level of education ranged from 5 to 27 years ( $M = 14.685$ ;  $SD = \pm 3.704$ ).

As for the cognitive reserve, assessed through the administration of the short Cognitive Reserve Index Questionnaire (sCRIq, Mondini et al., 2023), the average score was 112.8 ( $SD = \pm 18.886$ ) with values ranging from a minimum of 68 to a maximum of 169. *Table 3.1* summarizes the main descriptive stats of the sample involved.

The inclusion criteria for the sample recruitment are the following: being older than 18 years of age, being a native Italian speaker, having no diagnosis of neurological, psychiatric, or neurodegenerative disorders, and not showing any language impairment.

Among the 162 participants included in the normative sample, 56 were directly assessed by the present examiner, while the remaining data were collected by other examiners within the same research project, according to the same administration procedures.

All the people involved in the study participated voluntarily in the administration and were recruited through the examiner’s personal and extended social network.

	<i>Age</i>	<i>Education</i>	<i>CRI-Total</i>
<i>Mean</i>	47.12	14.69	112.8
<i>Mode</i>	30	13	101
<i>Median</i>	49	14.5	109
<i>Standard Deviation</i>	$\pm 18.425$	$\pm 3.704$	$\pm 18.886$

<i>Minimum</i>	18	5	68
<i>Maximum</i>	88	27	169
<i>Skewness</i>	0.254	-0.143	0.658
<i>Kurtosis</i>	2.03	3.36	3.17
<i>Shapiro-Wilk Test (p-value)</i>	<.001	<.001	<0.001

**Table 3.1:** Descriptive statistics of the Italian sample

Regarding the distribution of the main demographic variables taken into consideration, all age, education, and Cognitive Reserve Index (CRI) showed a non-normal distribution within the sample (Shapiro-Wilk test:  $p < .001$ , Image 3.1, Image 3.2, Image 3.3).

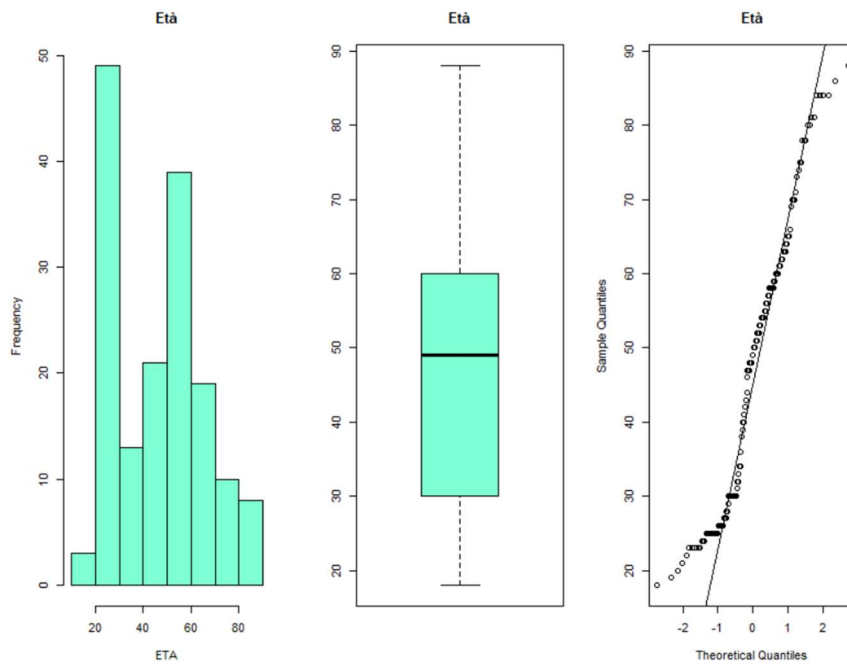


Image 3.1 age distribution in the Italian sample

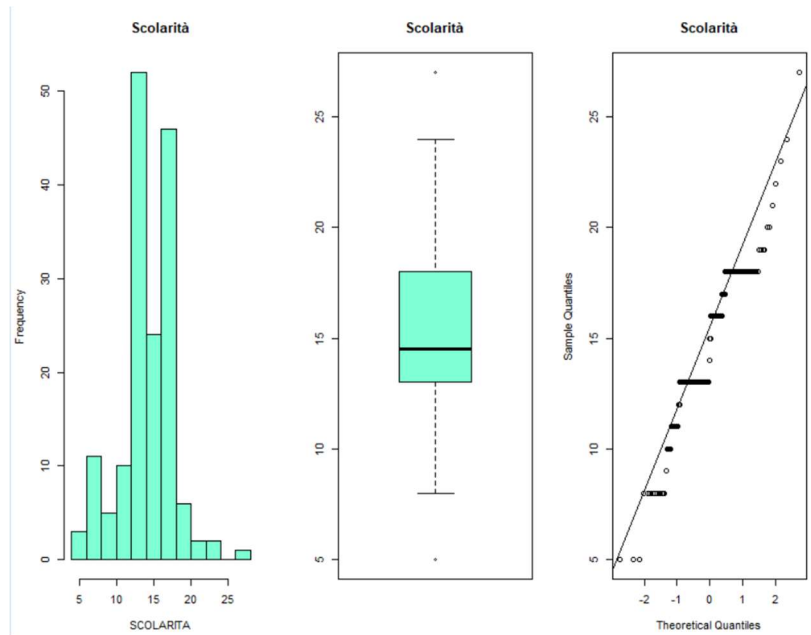


Image 3.2 Education distribution in the Italian sample

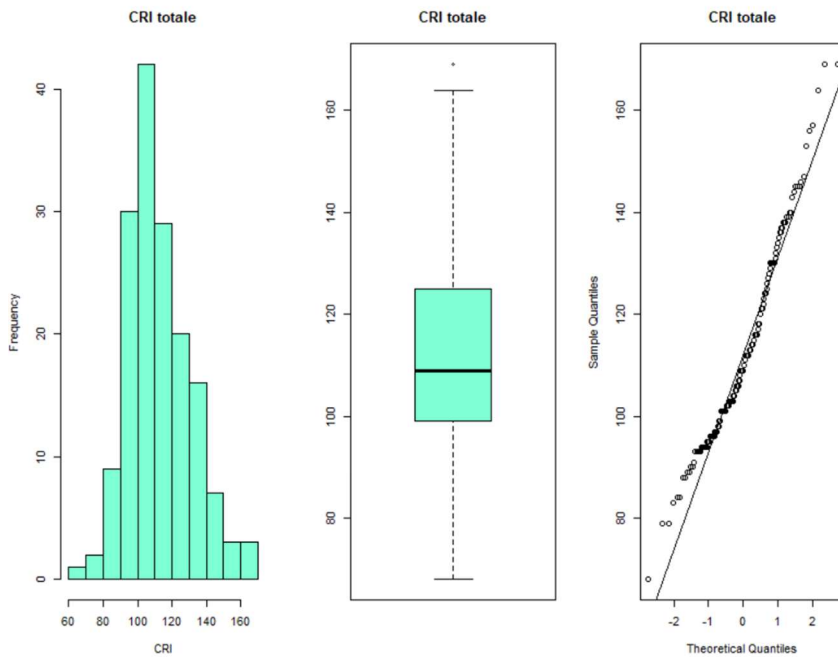


Image 3.3 CRI-Total distribution in the Italian Sample

### 3.1.2 Tools

The data collection process involved the administration of the *Non-Verbal Assessment (NoVA)* battery described in the previous chapter, the assessment of participants' cognitive reserve values (*CRI-Total*) through the administration of the short version of

the *Cognitive Reserve Index Questionnaire (s-CRIq)*, Mondini et al., 2023), and the weighted score obtained through the administration of the *Tele-Global Examination of Mental Health (tele-GEMS)*,  $\alpha$  et al., 2023). s-CRIq is available online at the following link: <https://www.cognitivereserveindex.org/NewEdition/calcolo.html>.

During a neuropsychological evaluation, it is very important to take cognitive reserve into account, as it represents an indicator of an individual's cognitive history. This index allows for a better understanding of the participant's overall cognitive functioning and can be of utmost importance during the interpretation of the results obtained from the neuropsychological tests. As a matter of fact, if we observe a person with a high cognitive reserve who obtains a score in the neuropsychological test below average, it could be helpful to delve deeper into the person's clinical picture, as the neurodegenerative process might already be at an advanced stage (Stern, 2012). Given the important role of cognitive reserve highlighted in the literature, this variable was included in the statistical analysis of the present study as a factor that could influence the participants' performance on the *NoVA Battery*. Specifically, the tool used to assess cognitive reserve was the *short Cognitive Reserve Index Questionnaire (s-CRIq)*, Mondini et al., 2023). This questionnaire collects information about cognitively stimulating experiences throughout an individual's adult life and quantifies them through standardized scores. The items included in the questionnaire focus on three main domains: education, working activities, and leisure time activities.

The global Cognitive Reserve Index (CRI) score allows the comparison of individuals with demographic characteristics, not only in terms of gender or age, but also in terms of years of education, occupational history, and engagement in stimulating activities carried out in leisure time, since the age of 18 (Nucci et al., 2012). The cognitive reserve index is composed of three domains that represent the three principal areas mentioned above: CRI-Education, CRI-Working Activity, and CRI-Leisure Time, which includes six cognitively stimulating activities (Mondini et al., 2016).

*CRI-Education* domain considers the number of years of formal education and additional training courses lasting at least six months. The number of years of schooling years automatically calculated based on the education level selected.

*CRI-Working Activities* domain evaluates the number of years the individual has spent working since the age of 18. The questionnaire allows the selection of up to five occupational activities.

*CRI-Leisure Time* domain includes six types of cognitively stimulating activities performed during the person's free time. Participants report the frequency of these activity (weekly or yearly), and the number of years they have been practicing since the age of 18. The six activities are: reading newspapers (not including social media), hobbies (such as artistic activities, gardening, puzzles, physical activities, etc.), reading books, attending exhibitions/concerts/conferences, traveling, and having children.

The concept of cognitive reserve was developed starting from the notion of brain reserve and from the observation that individuals with similar levels of brain pathology may show different cognitive outcomes. Cognitive reserve has therefore been conceptualized as the result of a person's cognitive history, including strategies, skills, and abilities developed through engagement in intellectually stimulating activities throughout life, such as education, occupational experiences, and leisure activities. Furthermore, at a neuronal level, cognitive reserve is considered the ability to cope with neuronal damage as a compensatory mechanism, due to aging or a brain accident, thanks to cognitive compensation strategies and the involvement of flexible alternative neuro-functional circuits to turn around the damaged synapses (Stern et al., 2014; 2019). These mechanisms help explain why individuals with similar levels of brain pathology may show different cognitive performances. For example, individuals who have engaged in cognitively stimulating activities throughout life may show higher cognitive reserve levels and may therefore experience a delayed onset of cognitive decline associated with aging or neurological conditions compared to individuals with lower cognitive reserve (Mondini et al., 2016).

### 3.1.3 Methods and Procedures

Participants were initially contacted by telephone and were informed about the research's aims, the approximate duration of the assessment, and the modalities of test administration. All individuals voluntarily agreed to participate in the study and did not receive any form of compensation.

Before the beginning of the assessment, all participants were provided with a written informed consent form containing detailed information regarding the study. The document was signed in a double copy: one retained by the participant, and the other by the examiner. The study was approved by the Ethics Committee of the University of Padova (protocol number 5233, dated 23/02/2023).

All participants completed the assessment following the same order of administration:

- 1) Short Cognitive Reserve Index questionnaire (s-CRIq, Mondini et al., 2023)
- 2) Non-Verbal Assessment Battery (NoVA)
- 3) Tele-Global Examination of Mental State (Tele-GEMS, Montemurro et al., 2023).

The s-CRIq and the NoVA battery were administered in person, either in the participants' homes or at the examiner's home. Although the setting was not a clinical environment, efforts were made to ensure the appropriate testing conditions. Assessments were conducted in quiet and well-lit rooms in order to minimize distractions and allow the correct administration of the test. To further reduce external interference, the presence of third parties during the testing session was not allowed. The Tele-GEMS assessment was administered remotely, through a phone call between the examiner and the participant, following the standardized administration procedure.

As concerns s-CRIq and NoVA battery, the data gathering took place in the participants' private homes or at the examiner's home, in a single session. Nevertheless, even if the setting was not characterized by a professional room, it was always possible to create a calm and healthy environment: a light and quiet room that allows the tests to be conducted without interferences and with the support of the correct tools. In order to minimize external interferences, it was forbidden for third parties to attend the testing session. The whole testing session lasted between 40 and 60 minutes.

Regarding Tele-GEMS, the administration took place remotely, through a phone call between the examiner and the examinee. The testing session lasted around 10 minutes.

### 3.1.4 Data Analysis

In this paragraph, the analyses conducted on the dataset collected from a sample of 162 healthy Italian participants are presented. Firstly, descriptive statistics were computed and analysed (Table 2). Subsequently, a multiple linear regression model was performed in order to determine whether and to what extent the main demographic variables significantly predict the global score obtained on the NoVA Battery. Additionally, further psychometric properties have been investigated, such as the internal consistency of the battery items and their construct validity (Table 3.2). All statistical analyses were conducted using the RStudio software environment (RStudio Team, 2020).

The global *NoVA* Battery score was calculated starting from the score of each subtest. Specifically, all raw scores were normalized so that each task contributed equally to the overall score. The procedure ensured that the resulting index of global cognitive functioning was balanced across the different cognitive domains assessed by the *NoVA* Battery. It was necessary to normalize the results as the tasks included in the battery differ in structure. For example, the *Logic Relations* task consists of four items, whereas the *Gesture Imitation* task consists of three items: without this normalization, some tasks would have weighted more on the global score than others. The normalized scores were then summed for each participant, and their means were calculated and multiplied by 100, resulting in a global cognitive functioning score, ranging from 0 to 100.

Descriptive statistical analyses were conducted by examining the distribution of the data for each subtest and for the global score using the *Shapiro-Wilk Test* (S-W). Concerning the global score, the distribution was found to be non-normal within the sample (S-W: < 0.001). Nonetheless, the score showed good variability, with a minimum value of 47.49 and a maximum value of 98.21.

A more detailed analysis of the distributions of the scores obtained in each subtest revealed that most tests showed non-normal distributions, generally characterized by negative skewness. The only exceptions in this study were the scores related to errors, omissions, and intrusions in memory tasks (both immediate and delayed), as well as the scores obtained in the Trail Making Tests (both *TMT-A* and *TMT-B*) for both

completion time and number of errors, as it shows positive skewness (Table 3.2). More precisely, for the *Colour Span Forward* and *Colour Span Backwards* tasks, the skewness values of the distributions, although negative, were relatively mild (-0.067 and -0.158, respectively). These values suggest that participants' performance on the tasks was relatively homogeneous, indicating that the level of difficulty of these tasks is well balanced for the sample taken into consideration. *Table 3.2* presents the descriptive statistics for the global score and the score of each subtest of the *NoVA Battery* obtained from the normative sample.

*Images 3.4* to *3.25* illustrate the distribution of the scores obtained in each task of the *NoVA Battery*.

	Mean	Mode	Median	SD	Min	Max	Skewness	Kurtosis
Total score	87.16		89.42	± 7.917	47.49	98.21	-1.717	7.68
Barrage	28.31	30	29	± 2.465	17	30	-2.142	7.997
TMT-A	16.44	13	15	± 7.959	4	51	1.256	5.075
Err. TMT-A	0.481	0	0	± 1.428	0	8	4.553	24.038
Corr. TMT-A	8.525	9	9	± 1.397	1	9	-4.537	24.36
TMT-B	32.87	27	28	± 16.424	9	101	1.693	6.408
Err. TMT-B	1	0	0	± 1.56	0	9	2.801	13.956
Corr. TMT-B	10.18	11	11	± 1.482	1	11	-3.089	16.551
SFMT	43.6	54	45	± 9.956	0	54	-0.992	4.273
Bi-dimensional drawing	1.864	2sdss	2	± 0.378	0	2	-2.792	10.475
Tri-dimensional drawing	1.84	2	2	± 0.415	0	2	-2.595	9.283
Movements imitation	2.827	3	3	± 0.395	1	3	-2.028	5.997
Motor sequence	2.512	3	3	± 0.797	0	3	-1.328	3.368
Incomplete figures	3.969	4	4	± 0.173	3	4	-5.425	30.431
Colour span forward	4.673	5	5	± 0.995	2	7	-0.067	3.118
Colour span backward	3.815	4	4	± 1.143	0	7	-0.158	3.907
Immediate memory	8.259	9	9	± 1.526	1	10	-1.199	5.369
Err. Immediate memory	1.698	1	1	± 1.495	0	9	1.274	5.794

Delayed memory	8.58	9	9	$\pm 1.264$	4	10	-0.892	3.612
Err. Delayed memory	1.407	1	1	$\pm 1.263$	0	6	0.903	3.652
Intrusions delay memory	1.302	0	1	$\pm 1.93$	0	11	2.449	10.52
Logic relations	3.377	4	4	$\pm 0.796$	0	4	-1.591	6.366
Interference	3.926	4	4	$\pm 0.41$	0	4	-7.024	59.072
Go/No-Go	3.907	4	4	$\pm 0.311$	2	4	-3.412	14.664

Table 3.2: Descriptive statistics for each subtest of the NoVA Battery f the Italian sample

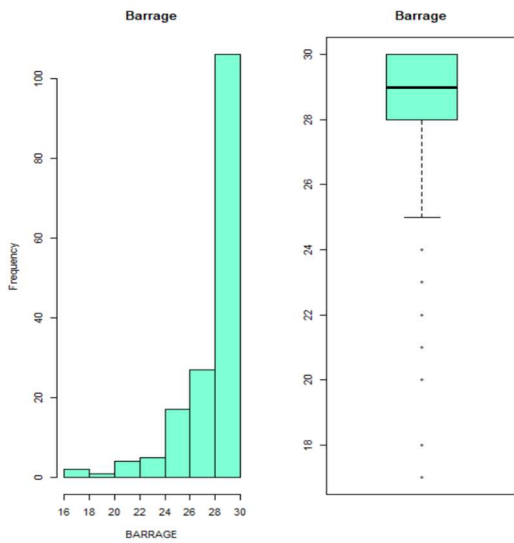


Image 3.4 Distribution of Flower Barrage scores

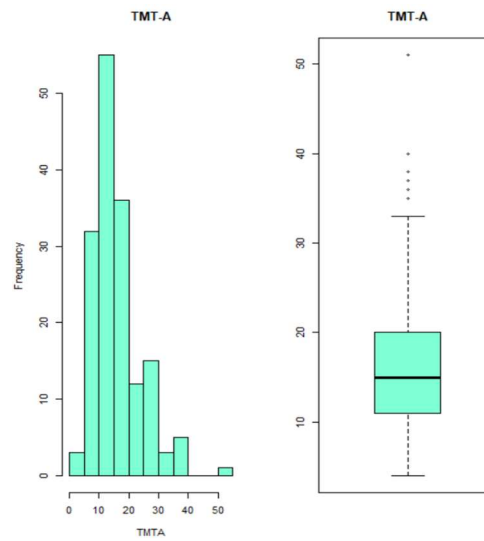


Image 3.5 Distribution of TMT-A scores

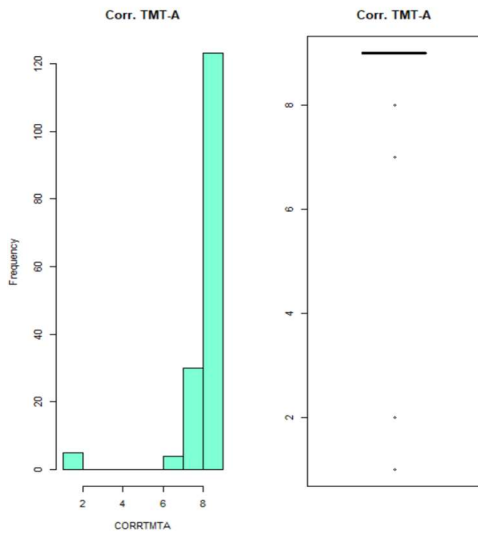


Image 3.6 Distribution of correct answers for TMT-A

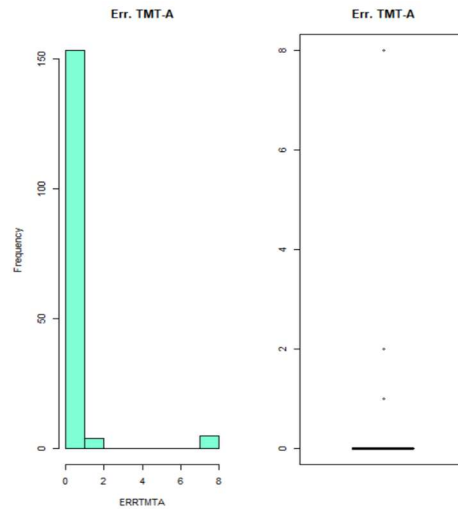


Image 3.7 Distribution of errors for TMT-A

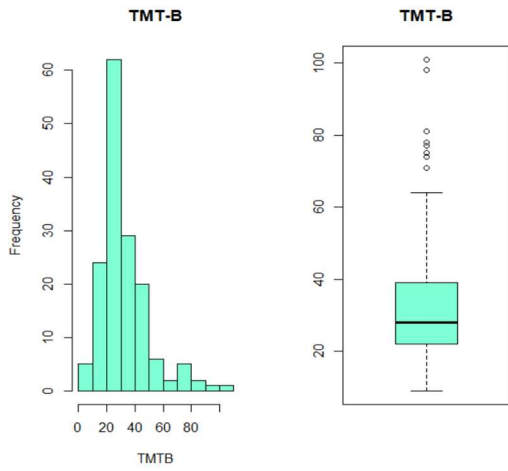


Image 3.8 Distribution of TMT-B scores

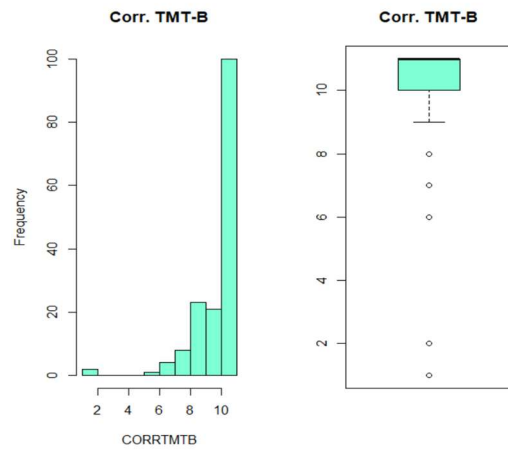


Image 3.9 Distribution of correct answers for TMT-B

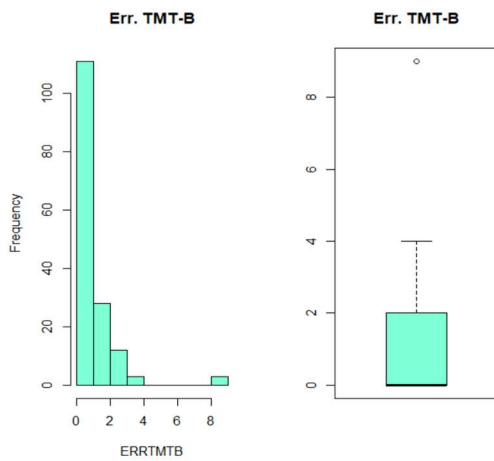


Image 3.10 Distribution of errors for TMT-B

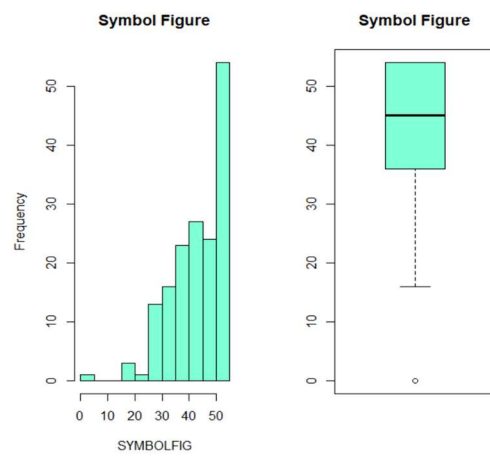


Image 3.11 Distribution of SFMT scores

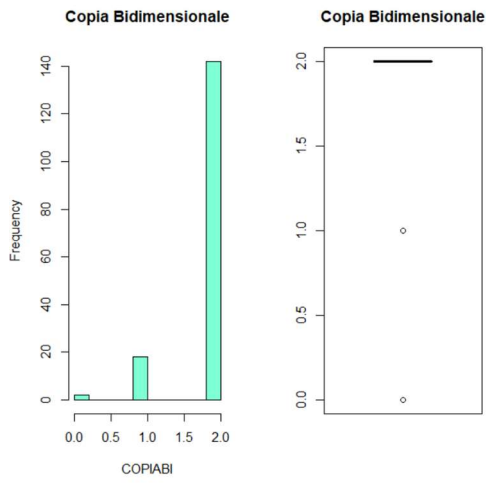


Image 3.12 Distribution of 2D drawing scores

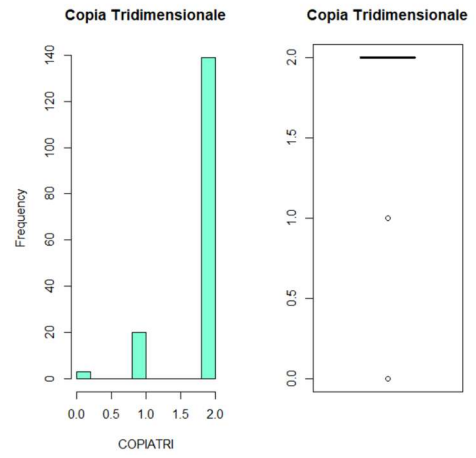


Image 3.13 Distribution of 3D drawing scores

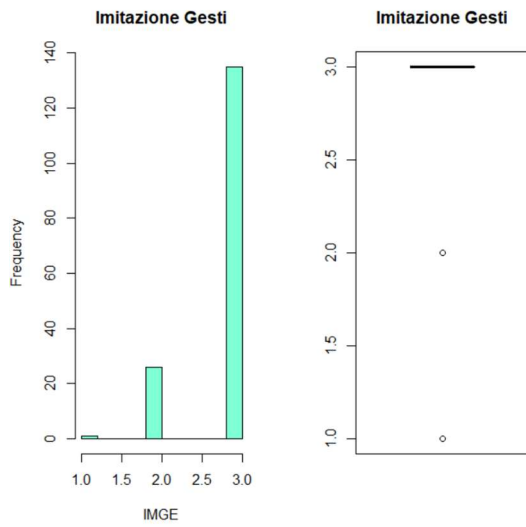


Image 3.14 Distribution of hand gesture imitations scores

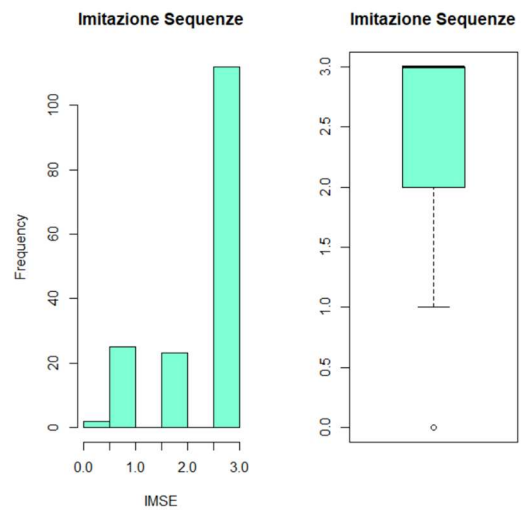


Image 3.15 Distribution of motor sequence imitations scores

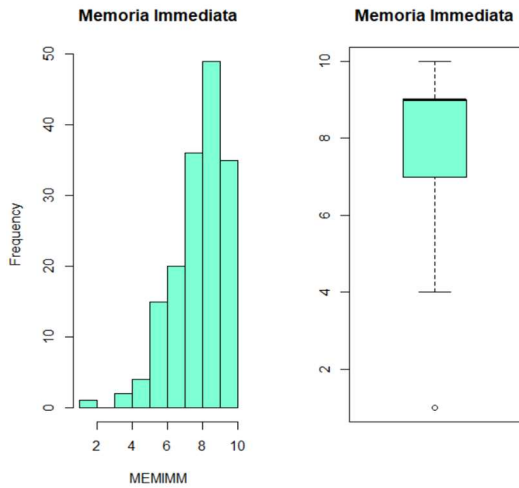


Image 3.16 Distribution of Immediate Recall scores

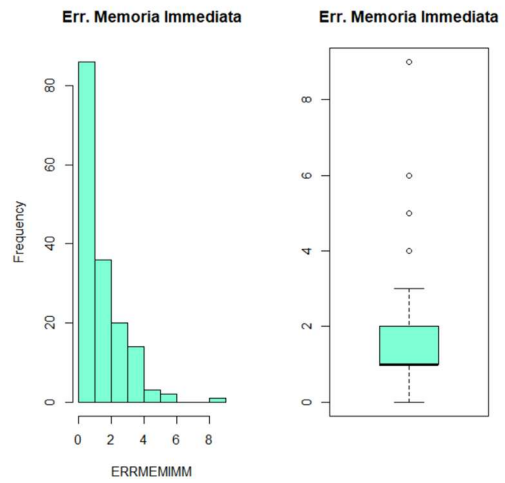


Image 3.17 Distribution of Immediate recall errors

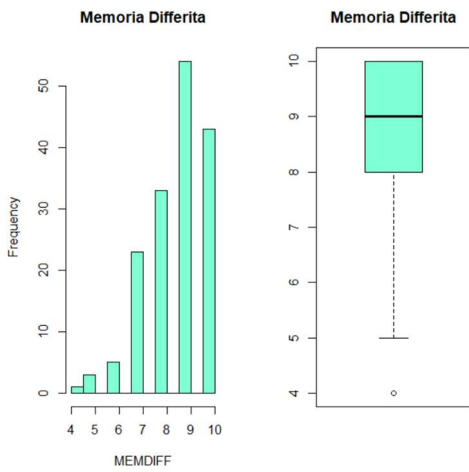


Image 3.18 Distribution of Delayed recall scores

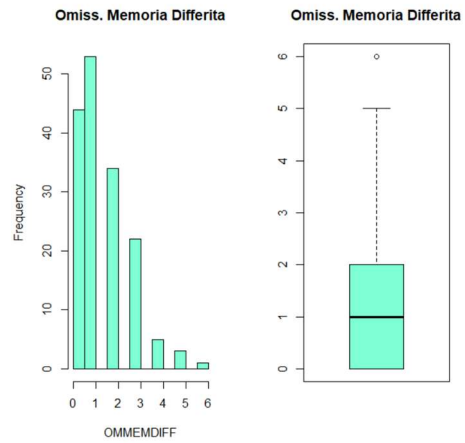


Image 3.19 Distribution of Delayed Recall omission

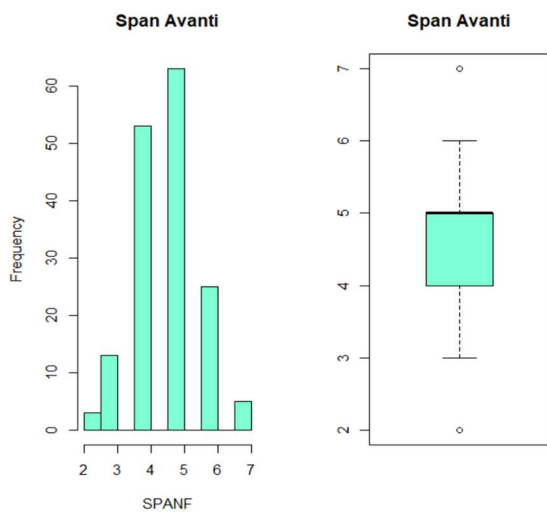


Image 3.20 Distribution of Colour Span F. scores

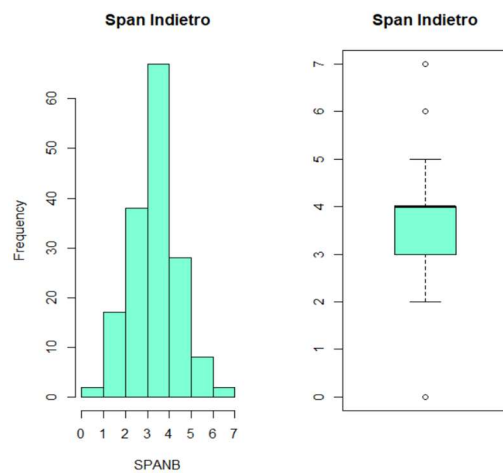


Image 3.21 Distribution of Colour Span B. scores

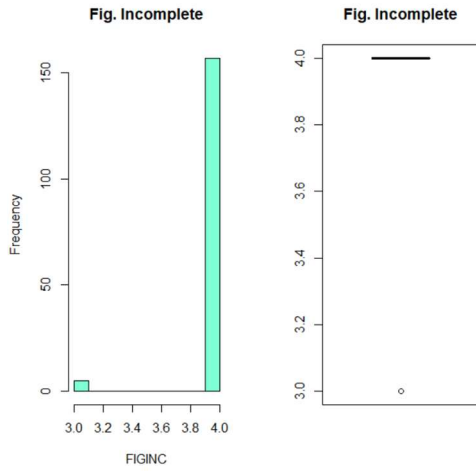


Image 3.22 Distribution of Incomplete Figures scores

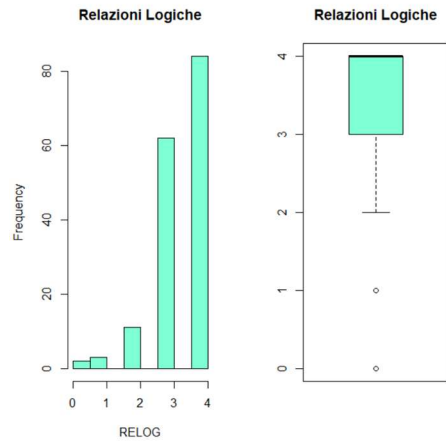


Image 3.23 Distribution of Logic Relations scores

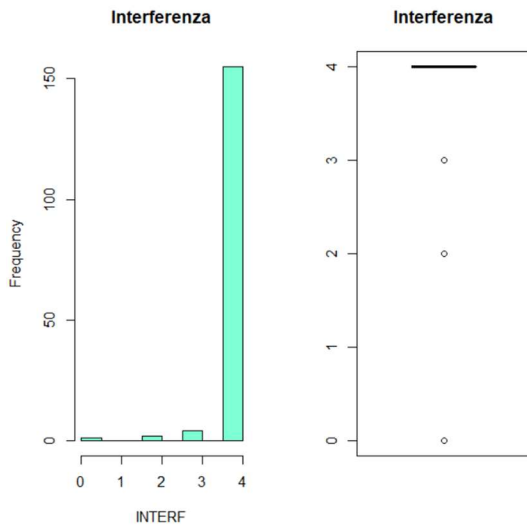


Image 3.24 Distribution of Interference scores

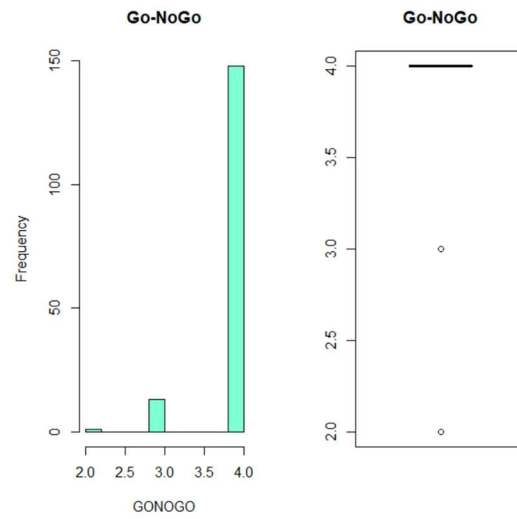


Image 3.25 Distribution of Go/No-Go scores

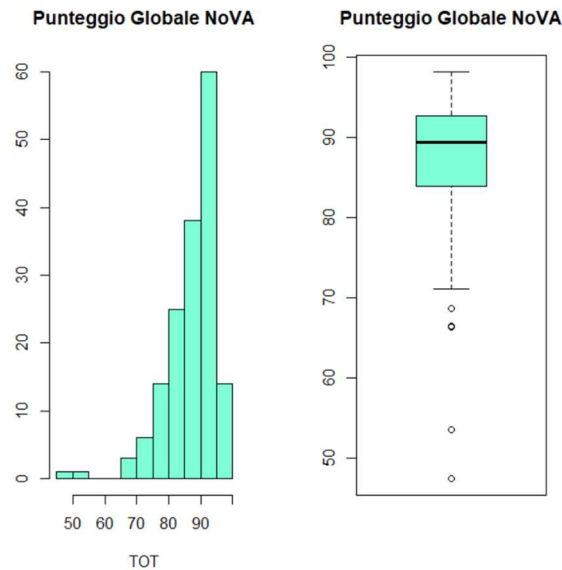


Image 3.26 Distribution of NoVA Global Scores for the Italian sample

Following the analysis of the descriptive statistics and the inspection of the score distribution, both for the global score and for each subtest, the psychometric properties of the instrument have been further investigated. Specifically, the internal consistency of the *NoVA* Battery was assessed by calculating *Cronbach's alpha coefficient*, considering the scores obtained in each subtest. The analysis showed that the instrument is characterized by a good level of internal consistency ( $\alpha = 0.59$ ; *standardized*  $\alpha = 0.81$ ), indicating a satisfactory degree of coherence among the tasks included in the battery.

Subsequently, in order to examine the effect of the main demographic variables on the prediction of the global *NoVA* score, multiple linear regression models were constructed. The model that best predicts the global score, based on the selected demographic variables, is presented below.

The regression model included age, years of education, and CRI as predictors of the global score. The model was found to be statistically significant ( $p < 0.001$ ) and showed an adjusted  $R^2$  of  $0.604$ . These results indicate that the model is able to explain a substantial portion of the variance in the global *NoVA* scores for the normative sample, accounting for slightly more than 60% of total variance.

All the variables included in the model emerged as significant predictors of the performance. In particular, age showed a significant negative association with the global score ( $\beta = -0.33$ ;  $p < .001$ ), suggesting that higher levels both in education and

in CRI predict better cognitive performance, especially education ( $\beta = 0.4$ ;  $p < 0.05$ ) determines higher values on the NoVA score when increasing. Table 3.3 presents the individual contribution of each predictor to the global NoVA score.

Predictors	Slope ( $\beta$ )	t-value	p-value
Intercept	86.407	35.267	<0.001***
Age	-0.33	-9.842	<0.001***
Education	0.4	2.531	0.012*
CRI	0.092	2.63	<0.01**

Table 3.3: Regression model with Global score as dependent variable and age and education as independent variables

$\beta$  = angular coefficient, indicates the slope of the regression line

p = p-value, probability value

R<sup>2</sup> = Multiple R-Squared, indicates the variance explained by the model

In neuropsychological testing, *cut-off* scores are thresholds used to classify an individual's cognitive performance relative to that of a reference group. These thresholds allow for determining whether a person's score falls within the normal range (thus within the distribution of the normative sample) or whether it lies below the expected range, typically below the 5<sup>th</sup> percentile, indicating the possibility of cognitive impairment.

The reference group – the normative sample – is usually composed of a large number of individuals whose characteristics are representative of the population of interest.

In the present study, *cut-off* scores were calculated based on the results obtained from the administration of the NoVA Battery to the entire sample of Italian participants. In particular, for each subtest, a linear regression model (Crawford and Garthwaite, 2006) was computed, including the main demographic variables (age, education, and CRI) as predictors. Based on these models, the predicted scores for each subtest were calculated.

Subsequently, the Mean and Standard Deviation (SD) of these predicted scores were derived. The *cut-off* for each test was then defined as the score that fell 1.5 standard

deviations below the mean of the predicted values for that specific task. The same procedure was applied to compute the *cut-off* for the global score (Table 3.4).

	Mean	Standard Deviation	Cut-off
Total score	87.159	± 7.917	74.02
Barrage	28.314	± 2.456	24.222
Corr. TMT-A	8.524	± 1.397	6.205
Corr. TMT-B	10.17	± 1.482	7.716
SFMT	43.604	± 9.956	27.083
2D drawing	1.864	± 0.378	1.236
3D drawing	1.839	± 0.415	1.149
Movements imitation	2.827	± 0.395	2.171
Motor sequence	2.512	± 0.797	1.188
Incomplete figures	3.969	± 0.173	3.681
Colour span forward	4.672	± 0.995	3.02
Colour span backward	3.814	± 1.143	1.917
Immediate memory	8.259	± 1.526	5.726
Delayed memory	8.58	± 1.264	6.481
Logic relations	3.376	± 0.796	2.055
Interference	3.925	± 0.41	3.244
Go/No-Go	3.907	± 0.311	3.39

Table 3.4: Mean, Standard Deviations, and cut-offs for each subtest of the NoVA Battery

Furthermore, descriptive statistics were computed for the total score and for each subtest of the *Tele-GEMS* administered to the normative sample (N = 162) of healthy Italian participants. The total score showed a mean of 91.24 (SD = ± 6.662), with a median of 92 and a mode of 98, indicating generally high performance within the sample. Score ranged from a minimum of 72.83 to a maximum of 100, suggesting

good variability despite an overall tendency toward higher scores. The distribution was characterized by negative skewness (-0.997), which suggests a concentration of scores in the upper range, and a slightly positive kurtosis (0.525), indicating a moderately peaked distribution. Taking a broader look at the single subtests, the *Orientation* subtest showed no variability (Mean = 4; Min = Max = 4), indicating that all participants achieved the maximum score. Overall, the results indicate that Tele-GEMS is characterized by high average performance and limited variability in most subtests.

Table 3.5 presents the descriptive statistics for the global score and the score of each subtest of the *Tele-GEMS* obtained from the normative sample.

	Mean	Mode	Median	SD	Min	Max	Skewness	Kurtosis
Tot Score	91.24	98	92	± 6.662	72.83	100	-0.997	0.525
Orientation	4	4	4	± 0	4	4	0	0
Imm. Memory	5.246	6	5	± 0.717	3	6	-0.962	-0.102
Del. Memory	4.632	4	5	± 0.966	3	6	-0.388	-1.007
Working Memory	4.719	5	5	± 0.437	3	5	-1.326	-0.276
Clock	3.895	4	4	± 0.392	3	4	-2.201	3.138
Naming	3.93	4	4	± 0.332	3	4	-1.7	0.975
Comprehension	0.929	1	1	± 1.014	0	1	-0.762	-1.578
Auditory Attention	7.702	8	8	± 0.437	6	8	-1.7	0.975
Verbal Fluency	14.86	13	15	± 2.573	9	24	0.113	-0.244
Metaphor	0.947	1	1	± 0.437	0	1	-1.7	0.975

Table 3.5: Descriptive statistics for each subtest of the *Tele-GEMS*

In order to explore the relationship between the two cognitive instruments, a correlation analysis was conducted between the global score of the NoVA Battery and the total score of the Tele-Gems, within the Italian sample. The results showed a moderate positive correlation ( $r = 0.528$ ), indicating that higher performance on the NoVA Battery is positively associated with higher performance on the Tele-GEMS. This

finding suggests that, although the two instruments differ in modality of administration and general structure, they capture partially overlapping aspects of global cognitive functioning.

### **3.2 Administration of NoVA Battery to Non-Native speakers**

The NoVA battery (Appendix) is designed as a cognitive assessment tool composed primarily of non-verbal tasks. This characteristic makes the instrument potentially useful for evaluating cognitive functioning in people with a language or communication impairment (for example, people with aphasia) or in situations where the assessment is conducted in a language different from the participant's native language. The tasks included in the battery rely on visual-spatial stimuli rather than alphanumeric ones. Furthermore, although the instructions are provided verbally by the examiner, they are matched with visual supports, including animated examples illustrating the task to be performed, ensuring adequate task comprehension. Moreover, the answer required from the examinee is non-verbal, reducing the linguistic demands of the assessment. The following section presents the results obtained by the administration of the NoVA battery to a sample of non-native speakers. Such in-depth analysis aims to test the "non-verbal" feature of the tool qualitatively in an environment characterized by a linguistic barrier, which differs from conditions involving language impairment but may still influence test performance.

#### **3.2.1 Sample, Materials, and Methods**

The sample taken under examination is composed of 20 non-native English-speaking people. The participants have been recruited through the examiner's personal contacts, including acquaintances and members of their extended social network. All people who took part in the study adhered voluntarily and without any kind of compensation.

The exclusion criteria considered for the data collection were the following: age of at least 18, being an English native speaker, and being diagnosed with a neurologic, neuropsychological, or psychiatric impairment. The countries of belonging of the people involved in the study are the following: Germany, Lithuania, the Netherlands, Slovenia, Turkey, and Ukraine. To all the participant the test was administered in English, and the language proficiency levels were variegated. *Table 3.6* describes the characteristics of the sample taken under examination.

	Age	Education	CRI-Total
Mean	35.86	16.86	114.6
Mode	24	18	103
Median	25	18	103
Standard Deviation	± 15.662	± 2.851	± 24.231
Minimum	24	8	92
Maximum	64	22	166
Skewness	0.746	-1.576	1.201
Kurtosis	1.69	6.082	2.828
Shapiro-Wilk Test (p-value)	<0.001	<0.001	<0.001

Table 3.6: Descriptive statistics of the non-native sample

Regarding the distribution of the main demographic variables taken into consideration, all age, education, and Cognitive Reserve Index (CRI) showed a non-normal distribution within the sample (Shapiro-Wilk test:  $p < .001$ , Image 3.27, Image 3.28, Image 3.29).

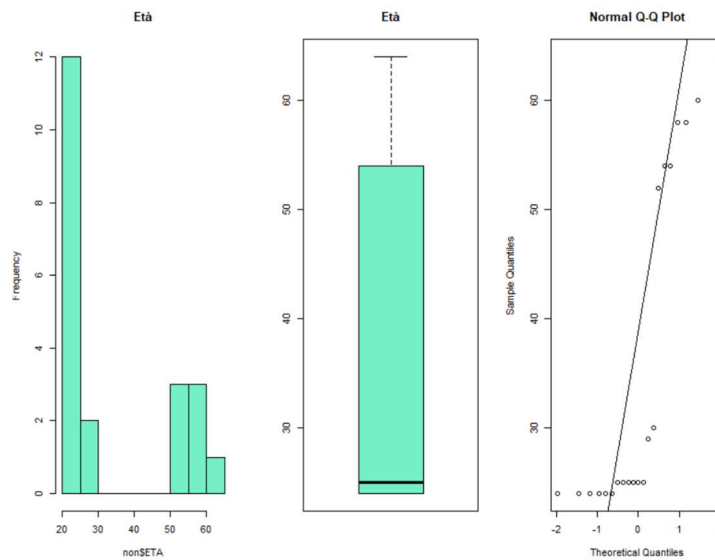


Image 3.27 age distribution in the Non-Native Sample

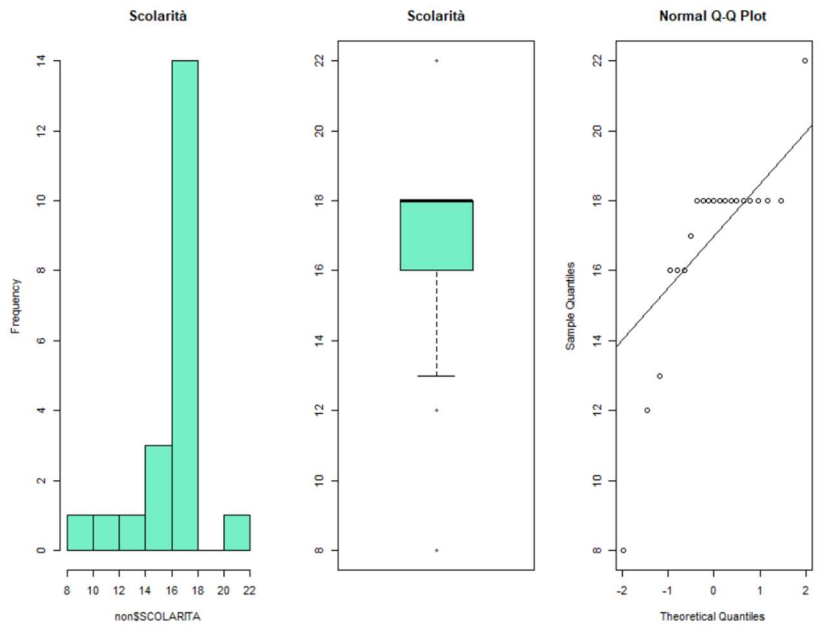


Image 3.28 education distribution in the Non-Native sample

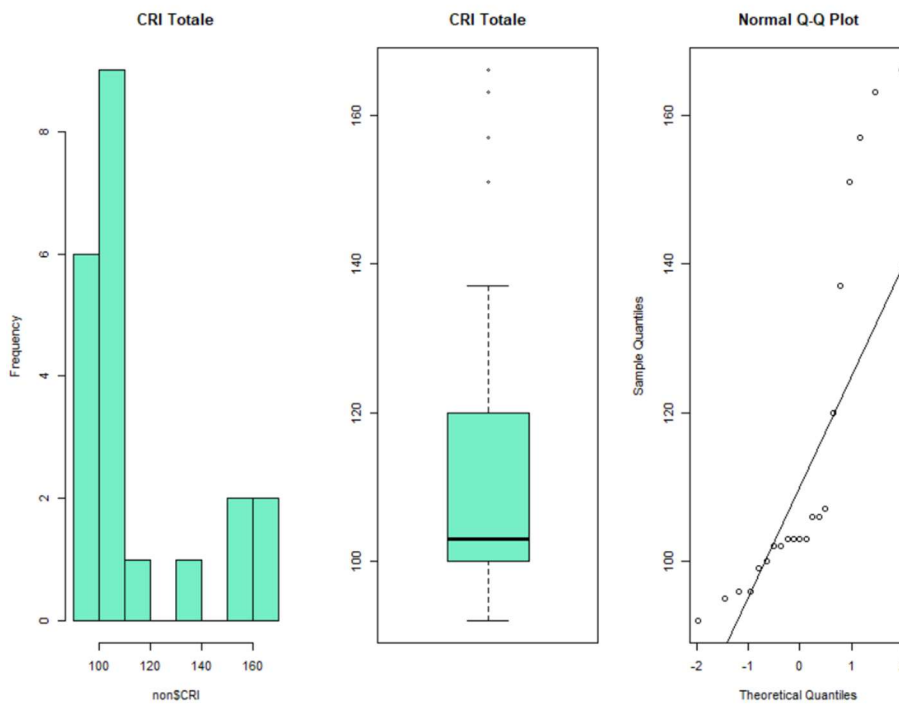


Image 3.29 CRI distribution in the Non-Native sample

As concerns s-CRIq and NoVA battery, the data gathering took place in the participants' private homes or at the examiner's home, in a single session. Nevertheless, even if the setting was not characterized by a professional room, it was always possible to create a calm and healthy environment: a light and quiet room that allows the tests to be conducted without interferences and with the support of the correct tools. In order to minimize external interferences, it was forbidden for third parties to attend the testing session. The whole testing session lasted between 40 and 60 minutes. The only available version of the informed consent form was in Italian. For this reason, the examiner ensured that participants were carefully informed about the study's aims and procedures, helping them understand the document. Subsequently, the completed and signed informed consent form was collected.

### 3.2.2 Results

Table 3.7 presents the descriptive statistics of the scores obtained by the Non-Native sample. Image 3.30 shows the distribution of the Total score for the Non-Native sample.

	Mean	Mode	Median	SD	Min	Max	Skewness	Kurtosis
Total score	92.88		93.76		85.44	98.33	-0.521	2.035
Barrage	28.9	30	29	± 1.225	26	30	-0.828	-0.164
TMT-A	14.57	17	15	± 4.158	6	25	0.319	1.405
Err. TMT-A	0.333	0	0	± 0.587	0	2	1.521	1.636
Corr. TMT-A	8.667	9	9	± 0.587	7	9	-1.521	1.636
TMT-B	28.95	27	27	± 8.692	15	49	0.534	0.266
Err. TMT-B	0.476	0	0	± 0.604	0	2	1.003	0.188
Corr. TMT-B	10.52	11	11	± 0.604	9	11	-1.003	0.188
SFMT	50.95	54	54	± 4.73	39	54	-1.309	0.511
Bi-dimensional drawing	2	2	2	0	2	2	0	0
Tri-dimensional drawing	1.905	2	2	± 0.307	1	2	-2.887	7.03
Movements imitation	2.952	3	3	± 0.223	2	3	-4.472	20
Motor sequence	2.857	3	3	± 0.366	2	3	-2.12	2.775
Incomplete figures	4	4	4	0	4	4	0	0
Colour span forward	4.523	5	5	± 0.825	3	6	-0.176	-0.212
Colour span backward	3.523	3	3	± 0.887	2	5	0.084	-0.525
Immediate memory	8.714	10	9	± 1.496	5	10	-1.002	0.335
Err. Immediate memory	1.286	0	1	± 1.496	0	5	1.002	0.335
Delayed memory	9.095	10	10	± 1.099	7	10	-0.636	-1.086

Err. Delayed memory	0.857	0	0	$\pm 1.071$	0	3	0.787	-0.723
Intrusions delay memory	0.523	0	0	$\pm 0.887$	0	3	1.592	1.854
Logic relations	3.667	4	4	$\pm 0.587$	2	4	-1.521	1.636
Interference	4	4	4	0	4	4	0	0
Go/No-Go	3.952	4	4	$\pm 0.223$	3	4	-4.47	20

Table 3.7 Descriptive statistics for each subtest of the NoVA Battery for the Non-Native sample

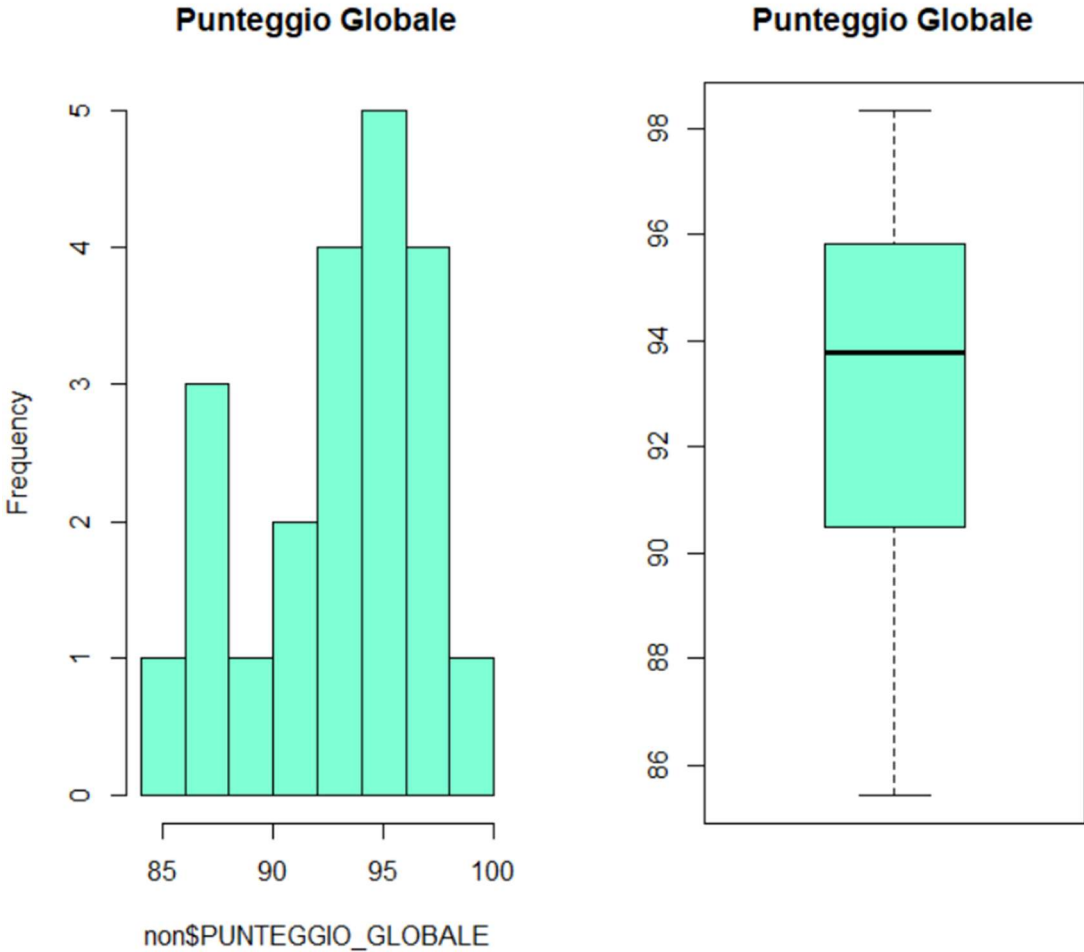


Image 2.30: Distribution of NoVA's Total Scores for the Non-Native sample

As in the case of the Italian normative sample, following the analysis of the descriptive statistics, three linear regression models were constructed in order to investigate the effect of the main demographic variables on the prediction of the global score

(Crawford and Garthwite). In the first model, Age, Education, and CRI were included as predictors, while the global score of the NoVA Battery was considered as the dependent variable.

Table 3.8 presents the Individual effect of each predictor on the global NoVA score.

predictors	Slope ( $\beta$ )	t-value	p-value
Intercept	99.181	22.854	<0.001***
Age	-0.207	-2.246	0.038*
Education	-0.021	-0.067	0.947
CRI	0.013	0.218	0.829

Table 3.8: Regression model with Global score as dependent variable and age, education, and CRI as independent variables

$\beta$  = angular coefficient, indicates the slope of the regression line

p = p-value, probability value

$R^2$  = Multiple R-Squared, indicates the variance explained by the model

Subsequently, in order to examine the effect of the main demographic variables on the prediction of the global NoVA score, multiple linear regression models were constructed. The model that best predicts the global score, based on the selected demographic variables, is presented below.

The regression model included age, years of education, and CRI as predictors of the global score. The model was found to be statistically significant ( $p < 0.001$ ) and showed an adjusted  $R^2$  of 0.516. These results indicate that the model is able to explain a substantial portion of the variance in the global NoVA scores for the Non-Native sample, accounting for slightly more than 50% of total variance.

The only variable included in the model that emerged as a significant predictor of the global score was age, as it showed a significant negative association with the global score ( $\beta = -0.207$ ;  $p = 0.038$ ). The fact that the  $\beta$  has a negative value means that the NoVA global score increases as the value of age decreases.

Since the CRIq (the short version too, as a direct consequence) was developed for an Italian population, it was decided to analyse the predictors without taking into consideration the CRI value.

It is important to consider that it is based on an extremely small sample size: this limitation may introduce significant bias and could potentially lead to misleading or inaccurate conclusions.

Predictors	Slope ( $\beta$ )	t-value	p-value
Intercept	99.131	23.505	<0.001***
Age	-0.189	-4.861	<0.001***
Education	0.031	0.146	0.885

Table 3.9: Regression model with Global score as dependent variable and age and education as independent variables

$\beta$  = angular coefficient, indicates the slope of the regression line

p = p-value, probability value

$R^2$  = Multiple R-Squared, indicates the variance explained by the model

This second regression model (Table 3.9) emerged as a significant predictor of the global score. The model was found to be statistically significant ( $p < 0.001$ ) and showed an adjusted  $R^2$  of 0.541. These results indicate that the model explains a portion of the variance in the NoVA global scores for the Non-Native sample, accounting for about 55% of the total variance. The variable included in the model that emerged as a strong predictor of the global score was, once again, age, as it showed a significant negative association with the global score ( $\beta = -0.189$ ;  $p < 0.001$ ). The fact that  $\beta$  has a negative value underlines that the global score of the test decreases as the value of age increases.

For purely statistical purposes, it was decided to examine how the variable age, considered independently, behaves within the linear regression model with the global score. The results are presented in the table 3.10 below.

Predictors	Slope ( $\beta$ )	t-value	p-value
Intercept	99.706	69.786	<0.001***
Age	-0.19	-5.198	<0.001***

Table 3.10: Regression model with Global score as dependent variable and age as independent variables

$\beta$  = angular coefficient, indicates the slope of the regression line

p = p-value, probability value

$R^2$  = Multiple R-Squared, indicates the variance explained by the model

This model was statistically significant ( $p < 0.001$ ) and had an adjusted  $R^2$  of 0.565. This result indicates that the model explains a good portion of variance in the NoVA global scores for the Non-Native sample, accounting for 56% of the total variance. The negative association with the global score ( $\beta = -0.19$ ;  $p < 0.001$ ) confirmed that increasing age is associated with lower performance on the global scores.

In addition to analyses conducted on the full normative and non-native sample, a further analysis was performed on two homogeneous groups. The purpose of this procedure was to reduce the demographic differences between the two groups and to allow a more controlled comparison of the NoVA score. Specifically, participants were selected from both Italian and Non-Native samples, in order to obtain two groups with similar values of age, education, and Cognitive Reserve Index (CRI). Images 3.31 and 2.32 show the distribution of global scores for the Italian and Non-Native subgroups, respectively.

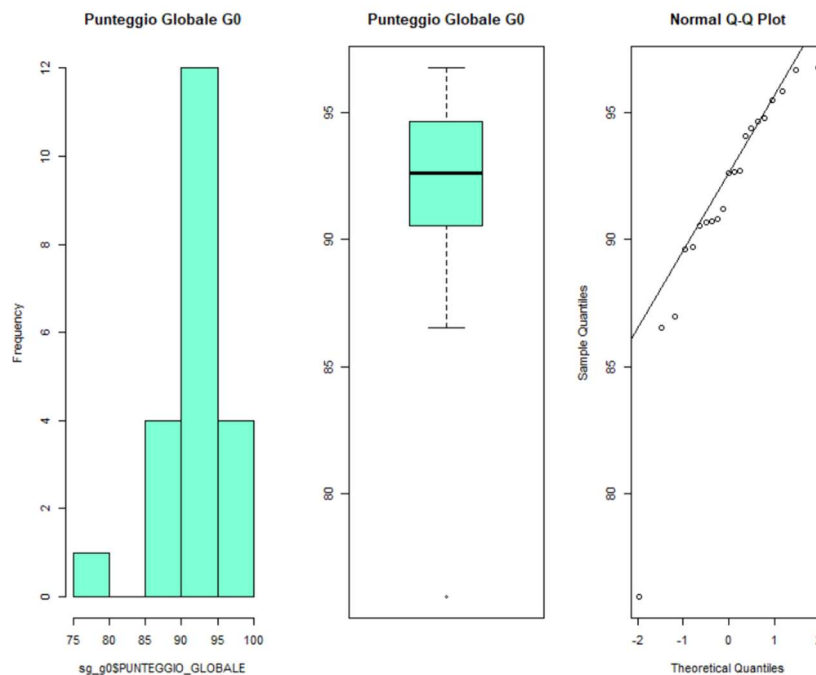


Image 3.31: Distribution of the NoVA scores of the Italian subgroup

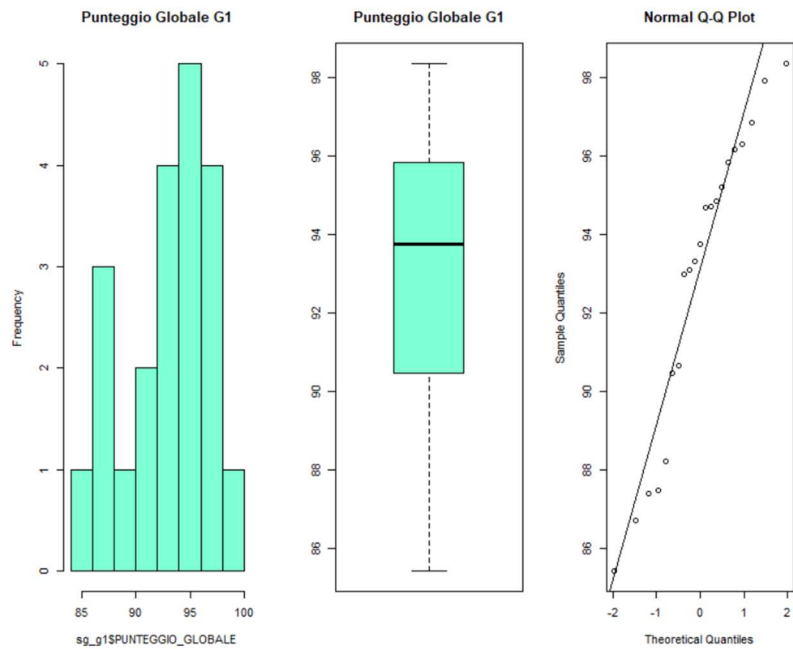


Image 3.32: distribution of the NoVA scores for the Non-Native subgroup

Before comparing the two subgroups, the distribution of the variables was assessed through the Shapiro-Wilk test: for the Italian subgroup, the NoVA score showed a non-normal distribution ( $p < 0.1$ ), whereas the Non-Native subgroup did not significantly differ from normality ( $p = 0.098$ ). The same procedure was applied to the matching variables: age showed a non-normal distribution for both subgroups (in both Italian and Non-Native subgroups,  $p < 0.001$ ), as well as for education (in both Italian and Non-Native subgroups,  $p < 0.001$ ), and for CRI (Italian subgroup,  $p < 0.01$ ; Non-Native subgroup,  $p < 0.001$ ). Since none of the variables showed a normal distribution, it was decided to carry out the comparison between the subgroups with the Wilcoxon rank-sum test. The results showed no significant differences between the two groups in the variable age ( $p = 0.817$ ), education ( $p = 0.723$ ), and CRI ( $p = 0.65$ ). Consequently, the two groups can be considered homogeneous for the variables used for the matching procedure.

The same test was performed to compare the global NoVA score between the two homogeneous subgroups. The analysis did not show a significant difference between the native and non-native groups ( $p = 0.33$ ). Overall, the comparison between the two homogeneous subgroups did not show significant differences in the NoVA global scores.

## **Discussions and conclusions**

The present study aimed to collect preliminary normative data for the NoVA Battery in a sample of healthy Italian participants and to investigate some of its psychometric properties. Concurrently, the study explored the administration of the aforementioned battery in a smaller sample of non-native speakers in order to examine whether a tool characterized by reduced linguistic demands could be supportive in situations of cognitive assessment with people from different linguistic and cultural backgrounds. In this sense, NoVA Battery was examined as a new multidomain neuropsychological instrument, as well as a possible response to the central issue represented by the role of language as a methodological variable in cognitive assessment.

Firstly, in the normative Italian sample, composed of 162 healthy participants, performance on the NoVA Battery was generally high, as expected in a group of healthy participants without neurological or psychiatric conditions. This pattern is consistent with the intended use of the battery, as no significant difficulties in task execution would be expected in a cognitively intact population.

The tendency toward higher scores was reflected in several subtests. More specifically, Interference, Go/No-Go, Incomplete Figure, and Movements Imitation tasks showed high mean values. Similarly, in Bi-Dimensional and Tri-Dimensional drawing tasks, the means of the sample were very close to the maximum score. These values suggest that a large part of the sample completed these tasks with relative ease. From a psychometric point of view, this pattern suggests the presence of a ceiling effect in some parts of the battery, which does not necessarily reduce the overall value of the instrument, but may limit the interpretation of performance in individuals at the upper end of performance.

However, Colour Span Forward and Colour Span Backward tasks appeared more balanced in terms of difficulty, with lower mean scores and a less marked asymmetry than other subtests. This suggests that these tasks allow a wider variability of performance, even in a healthy sample. Although Colour Span tasks were designed to reduce verbal mediation by replacing digits with colours, they still assess short-term memory and working memory, functions that might be supported by “internal” verbal strategies.

On this matter, the present data are in line with the broader assumption discussed in

the theoretical framework: the influence of language can be reduced, but not completely removed, especially in tasks involving memory, attention, and executive control (Rosselli & Ardila, 2003; Ardila, 2005).

An important methodological element of this study refers to the way the global NoVA score was calculated. The tasks included in the battery have different formats and different score ranges: to avoid this problem, raw scores were normalized in order to ensure that each subtest contributed equally to the score. In fact, tasks such as the Symbol Figure Modality Test or the Flower Barrage task have more items and wider score ranges; consequently, they would have had a stronger impact on the final index compared to other tasks characterized by fewer items. The use of a normalized score makes the global measure more balanced and more consistent for giving an overview of cognitive functioning across different non-verbal domains, rather than privileging one task.

Another relevant aspect that emerged from the analysis concerns the internal consistency of the battery. Cronbach's alpha value indicated a satisfactory level of coherence among the tasks included in the NoVA Battery. This value is particularly relevant considering that the NoVA Battery was not designed to investigate isolated functions, but to explore different cognitive domains. In this sense, the standardised alpha suggests that the battery maintains a good balance between multidomain structure and psychometric coherence.

The regression analysis conducted on the normative sample also offers important information for interpreting the performance on the NoVA Battery. The inclusion of age, education, and cognitive reserve as predictors highlights the extent to which individual differences in cognitive functioning are influenced by well-established demographic variables. In particular, the observed pattern suggests that higher levels of education and cognitive reserve are associated with better performance, whereas increasing age is associated with a decline in performance. These findings are consistent with the literature and with the rationale of the present thesis. Higher levels of education and of cognitive reserve are generally associated with better cognitive outcomes, whereas age (especially in functions such as processing speed, selective attention, and working memory) is influenced negatively by age (Stern, 2009; Lezak et al., 2012).

Another relevant aspect of the present work consists of the relationship between the NoVA Battery and the Tele-GEMS. As clarified in the methodological paragraph in Chapter 3, Tele-GEMS was not the focus of the study but was included as an already established neuropsychological screening tool, used to support the construct validity of NoVA. The association observed between the two instruments suggests that performance on the NoVA Battery reflects, at least partially, global cognitive functioning as measured by a widely established screening tool. At the same time, the relationship is not so strong as to indicate full overlap between the two measures. This is likely due to their different structural characteristics: while Tele-GEMS is shorter, and a fully verbally mediated instrument administered remotely, NoVA relies on in-person administration and is based on visually mediated and low-language tasks. This partial convergence supports the idea that NoVA captures core aspects of cognitive functioning, while also providing complementary information that is less influenced by linguistic mediation. In this sense, the battery might represent a useful addition to existing tools, particularly in contexts where language could act as a confounding factor.

A particularly relevant aspect of the study concerns the administration of the NoVA Battery to a Non-Native sample that included 20 participants. This section has an exploratory nature but is directly connected to the rationale of the thesis. As discussed in the 1<sup>st</sup> Chapter, language might influence neuropsychological assessment, not only concerning verbal production, but also for instruction comprehension, rule maintenance, and adaptation to the testing environment (Baddeley, 2003; Ardila, 2005; Hodges et al, 2005). The results obtained from the non-native sample provide preliminary support for this rationale. Participants were generally able to complete the tasks successfully, suggesting that the structure of the battery allows effective administration even in the presence of a linguistic mismatch between the test language and the participant's native language. However, these results suggest that the battery was administrable to the sample, and that the participants were generally able to complete the task successfully, even in the presence of a linguistic difference between the test administration language and the participant's native language.

These findings need to be cautiously interpreted. The difference in sample size between the two samples is substantial: the normative sample included 162 participants, while the Non-Native sample included only 20 participants. This sample-

size difference has important consequences for the stability and generalizability of the results. In addition, the Non-Native group of participants was more heterogeneous in terms of origin, culture, and English proficiency, which was the language of administration.

In order to help to better frame the comparison between Italian and Non-Native participants, two more heterogeneous groups were created. In addition to the analyses performed on the full samples, participants were selected from the two original groups for the purpose of obtaining two groups with similar values of age, education, and CRI, thus allowing a controlled comparison between the two linguistic groups. The comparison between the matched groups did not reveal substantial differences in performance on the global NoVA score. This finding suggests that, when demographic variables are controlled, linguistic background may have a reduced impact on performance when using a battery characterised by limited linguistic demands. These findings must not be interpreted as evidence of complete language neutrality, but rather as further indications that the NoVA Battery might reduce linguistic confounding associated with traditional neuropsychological assessment tools.

The regression analysis conducted on the Non-Native sample presented a pattern that differed from the Italian group: age emerged as the only consistent predictor of performance, whereas education and cognitive reserve did not show any significant contribution.

However, these results suggest that age plays an important role for the Non-Native sample; however, caution is needed here as well: more plausibly, the absence of a significant effect from the variables education and CRI is reflected in the limited sample size and the difficulty in evaluating the proper Cognitive Reserve Index in participants from different cultural backgrounds, since the CRIq was developed in relation to an Italian population, its use in a more heterogeneous population might not reflect the cognitive reserve in a fully comparable way. These findings obtained from the Non-Native group support the idea that reduced verbal demands are not necessarily a culture-free battery: NoVA Battery appears to reduce the need for verbal production and to limit part of the linguistic load usually involved in neuropsychological tools. Anyway, the performance might be influenced by other variables, such as education level, cultural background, familiarity with testing situations, and problem-solving strategies used in everyday tasks. In these contexts, cognitive scores should always

be read in relation not only to possible impairments or diagnoses, but also to demographic, educational, and sociocultural variables that may shape the testing situation.

Several limitations need to be acknowledged:

- Even if the normative sample might be considered adequate for a preliminary psychometric investigation, it is not broad enough to define stable norms for general use.
- Some subtests appear to show ceiling effects.
- The Assessments were conducted in non-clinical settings, such as participants' homes or the examiner's home, although efforts were made to maintain the correct testing conditions.
- The Non-Native sample was much smaller than the normative one, and highly heterogeneous. Furthermore, no formal measure of English proficiency was included.
- The study focused mainly on preliminary normative data, construct validity, and demographic predictors. Other relevant psychometric measures, such as test-retest reliability and inter-rater reliability, were not explored.

In conclusion, NoVA Battery appears to be a promising tool for the evaluation of cognitive functioning, especially in situations where language may interfere with standard testing conditions. The results of the present study suggest that reducing linguistic demands can contribute to a more accessible evaluation of cognitive performance across individuals with different linguistic backgrounds. Further studies on larger samples and more heterogeneous linguistic groups will be necessary to better define the psychometric properties and the potential clinical applicability of the battery.



## REFERENCES

- Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology review*, 15(4), 185-195.
- Ardila, A. (2007). The impact of culture on neuropsychological test performance. *International handbook of cross-cultural neuropsychology*, 23, 44.
- Ardila, A. (2021). Cross-cultural differences in cognition and learning. *The SAGE handbook of evolutionary psychology: Foundations of evolutionary psychology*, 420-435.
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of communication disorders*, 36(3), 189-208.
- Bennett, T. L. (2001). Neuropsychological evaluation in rehabilitation planning and evaluation of functional skills. *Archives of Clinical Neuropsychology*, 16(3), 237-253.
- Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of clinical neuropsychology*, 22(3), 355-365.
- Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of child psychology and psychiatry*, 46(3), 317-326.
- Brown, L. (2003). Test of nonverbal intelligence: A language-free measure of cognitive ability. In *Handbook of nonverbal assessment* (pp. 191-221). Boston, MA: Springer US.
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1990). *Test of Nonverbal Intelligence: A language-free measure of cognitive ability*. Austin, TX: Pro-ed.
- Butler, C. R., Brambati, S. M., Miller, B. L., & Gorno-Tempini, M. L. (2009). The neural correlates of verbal and nonverbal semantic processing deficits in neurodegenerative disease. *Cognitive and behavioral neurology*, 22(2), 73-80.

- Carone, D. A. (2007). E. Strauss, EMS Sherman, & O. Spreen, A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary: A Review of:“, Oxford University Press, New York, 2006.”.
- Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology review*, 13(4), 181-197.
- Christy, E. M., & Friedman, R. B. (2005). Using non-verbal tests to measure cognitive ability in patients with aphasia: A comparison of the RCPM and the TONI. *Brain and Language*, 95(1), 195-196.
- Cormier, D. C., Bulut, O., McGrew, K. S., & Kennedy, K. (2022). Linguistic influences on cognitive test performance: Examinee characteristics are more important than test characteristics. *Journal of Intelligence*, 10(1), 8.
- Cubillos, C., & Rienzo, A. (2023). Digital cognitive assessment tests for older adults: systematic literature review. *JMIR mental health*, 10(1), e47487.
- Daga, M., Mohanty, P., Krishna, R., & RM, S. P. (2025). Statistical Validation in Cultural Adaptations of Cognitive Tests: A Multi-Regional Systematic Review. *arXiv preprint arXiv:2504.13495*.
- Daugherty, J. C., Puente, A. E., Fasfous, A. F., Hidalgo-Ruzzante, N., & Pérez-Garcia, M. (2017). Diagnostic mistakes of culturally diverse individuals when using North American neuropsychological tests. *Applied Neuropsychology: Adult*, 24(1), 16-22.
- Davis, J. L., & Matthews, R. N. (2010). NEPSY-II Review: Korkman, M., Kirk, U., & Kemp, S.(2007). NEPSY—Second Edition (NEPSY-II). San Antonio, TX: Harcourt Assessment. *Journal of Psychoeducational Assessment*, 28(2), 175-182.
- De Meuse, K. P. (1987). A review of the effects of non-verbal cues on the performance appraisal process. *Journal of Occupational Psychology*, 60(3), 207-226.

- Diaz-Asper, C. M., Schretlen, D. J., & Pearlson, G. D. (2004). How well does IQ predict neuropsychological test performance in normal adults?. *Journal of the International Neuropsychological Society*, 10(1), 82-90.
- Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. F. A. B. (2000). The FAB: a frontal assessment battery at bedside. *Neurology*, 55(11), 1621-1626.
- Folstein, M. F., Robins, L. N., & Helzer, J. E. (1983). The mini-mental state examination. *Archives of general psychiatry*, 40(7), 812-812.
- Franzen, S., European Consortium on Cross-Cultural Neuropsychology (ECCroN), Watermeyer, T. J., Pomati, S., Papma, J. M., Nielsen, T. R., ... & Bekkhus-Wetterberg, P. (2022). Cross-cultural neuropsychological assessment in Europe: Position statement of the European consortium on cross-cultural neuropsychology (ECCroN). *The Clinical Neuropsychologist*, 36(3), 546-557.
- Fratti, S., Bowden, S. C., & Pino, O. (2011). Diagnostic memory assessment in Italian-born Australians. *International psychogeriatrics*, 23(7), 1133-1143.
- Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007). Language of administration and neuropsychological test performance in neurologically intact Hispanic American bilingual adults. *Archives of Clinical Neuropsychology*, 22(8), 991-1001.
- Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The bells test: a quantitative and qualitative test for visual neglect. *International journal of clinical neuropsychology*, 11(2), 49-54.
- Gkintoni, E., & Nikolaou, G. (2024). The cross-cultural validation of neuropsychological assessments and their clinical applications in cognitive behavioral therapy: A scoping analysis. *International Journal of Environmental Research and Public Health*, 21(8), 1110.
- Goodman, Z. T., Llabre, M. M., González, H. M., Lamar, M., Gallo, L. C., Tarraf, W., ... & Bainter, S. A. (2021). Testing measurement equivalence of neurocognitive assessments across language in the Hispanic Community Health Study/Study of Latinos. *Neuropsychology*, 35(4), 423.

- Gullifer, J. W., Koussaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., ... & Titone, D. (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, 42(2), 245-278.
- Harris, J. G., Tulskey, D. S., & Schultheis, M. T. (2003). Assessment of the non-native English speaker: Assimilating history and research findings to guide clinical practice. In *Clinical interpretation of the WAIS-III and WMS-III* (pp. 343-390). Academic Press.
- Harvey P. D. (2012). Clinical applications of neuropsychological assessment. *Dialogues in Clinical Neuroscience*, 14(1), 91–99. <https://doi.org/10.31887/DCNS.2012.14.1/pharvey>
- Hodges, J. R. (2017). *Cognitive assessment for clinicians*. Oxford University Press.
- Kassir, R., Roussel, M., Abboud, H., & Godefroy, O. (2025). Screening for cognitive impairment in bilinguals: What is the influence of the language of assessment?. *Canadian Journal of Neurological Sciences*, 52(2), 263-268.
- Kandiah, N., Zhang, A., Bautista, D. C., Silva, E., Ting, S. K. S., Ng, A., & Assam, P. (2016). Early detection of dementia in multilingual populations: Visual Cognitive Assessment Test (VCAT). *Journal of Neurology, Neurosurgery & Psychiatry*, 87(2), 156-160.
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY - Second Edition (NEPSY-II)*. Harcourt Assessment, San Antonio, TX.
- Lehman Blake, M., Ott, S., Villanyi, E., Kazhuro, K., & Schatz, P. (2015). Influence of language of administration on ImPACT performance by Bilingual Spanish–english college students. *Archives of Clinical Neuropsychology*, 30(4), 302-309.
- Lev-Ari, S. (2015). Comprehending non-native speakers: Theory and evidence for adjustment in manner of processing. *Frontiers in psychology*, 5, 1546.
- Lezak, M. D. (2004). *Neuropsychological assessment*. Oxford University Press, USA.
- Li, C., Sheng, Z., Cohen, T., & Pakhomov, S. V. (2025, May). "Is There Anything Else?": Examining Administrator Influence on Linguistic Features from the

Cookie Theft Picture Description Cognitive Test. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics* (pp. 91-103).

- Lim, L., Ng, T. P., Ong, A. P., Tan, M. P., Cenina, A. R., Gao, Q., ... & Kandiah, N. (2018). A novel language-neutral Visual Cognitive Assessment Test (VCAT): validation in four Southeast Asian countries. *Alzheimer's research & therapy, 10*(1), 6.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. *Handbook of second language acquisition*.
- Luque, A., Koronkiewicz, B., Issa, B., Faretta-Stutenberg, M., & Bowden, H. W. (2025). Ecological validity and inclusivity in heritage bilingualism research: Examining objective and subjective Spanish proficiency assessments and language experience factors. *Frontiers in Language Sciences, 3*, 1400587.
- Mancuso, V., Sarcinella, E. D., Bruni, F., Arlati, S., Di Santo, S. G., Cavallo, M., ... & Pedrolì, E. (2024). Systematic review of memory assessment in virtual reality: evaluating convergent and divergent validity with traditional neuropsychological measures. *Frontiers in human neuroscience, 18*, 1380575.
- McCallum, R. S. (Ed.). (2003). *Handbook of nonverbal assessment* (Vol. 30). New York: Kluwer Academic/Plenum Publishers.
- McLaughlin, B., & Nayak, N. (1989). Processing a New Language: Does Knowing Other Languages Make. *Interlingual processes, 1*, 5.
- Mondini, S., Mapelli, D., & Arcara, G. (2016). *Semeiotica e diagnosi neuropsicologica. Metodologia per la valutazione*. Carocci
- Mondini, S., & Mapelli, D. (Eds.). (2022). *Esame neuropsicologico breve 3. Con app. Con Contenuto digitale per download e accesso on line*. Cortina Raffaello.
- Montemurro, S., Mondini, S., Pucci, V., Durante, G., Riccardi, A., Maffezzini, S., ... & Arcara, G. (2023). Tele-Global Examination of Mental State (Tele-GEMS): an open tool for the remote neuropsychological screening. *Neurological Sciences, 44*(10), 3499-3508.

- Nardone, D. A., Johnson, G. K., Faryna, A., Coulehan, J. L., & Parrino, T. A. (1992). A model for the diagnostic medical interview: nonverbal, verbal, and cognitive assessments. *Journal of general internal medicine*, 7(4), 437-442.
- Ng, K. P., Chiew, H. J., Lim, L., Rosa-Neto, P., Kandiah, N., & Gauthier, S. (2018). The influence of language and culture on cognitive assessment tools in the diagnosis of early cognitive impairment and dementia. *Expert Review of Neurotherapeutics*, 18(11), 859-869.
- Nielsen, T. R. (2022). Cognitive assessment in culturally, linguistically, and educationally diverse older populations in Europe. *American Journal of Alzheimer's Disease & Other Dementias*®, 37, 15333175221117006.
- Norman, M. A., Evans, J. D., Miller, W. S., & Heaton, R. K. (2000). Demographically corrected norms for the California verbal learning test. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 80-94.
- Norman, M. A., Moore, D. J., Taylor, M., Franklin Jr, D., Cysique, L., Ake, C., ... & Hnrc Group. (2011). Demographically corrected norms for African Americans and Caucasians on the hopkins verbal learning test–revised, brief visuospatial memory test–revised, stroop color and word test, and wisconsin card sorting test 64-card version. *Journal of clinical and experimental neuropsychology*, 33(7), 793-804.
- Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive Reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging clinical and experimental research*, 24(3), 218-226.
- Razani, J., Burciaga, J., Madore, M., & Wong, J. (2007). Effects of acculturation on tests of attention and information processing in an ethnically diverse group. *Archives of Clinical Neuropsychology*, 22(3), 333-341.
- Reinhard, M. A., & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology*, 44(3), 477-488.
- Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and motor skills*, 8(3), 271-276.

- Rosselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and cognition*, 52(3), 326-333.
- Shaughnessy, L. W., & Weintraub, S. (2025). The role of neuropsychological assessment in the evaluation of patients with cognitive-behavioral change due to suspected Alzheimer's disease and other causes of cognitive impairment and dementia. *Alzheimer's & Dementia*, 21(1), e14363.
- Smith, A. (1973). *Symbol digit modalities test*.
- Sisco, S., Gross, A. L., Shih, R. A., Sachs, B. C., Glymour, M. M., Bangen, K. J., ... & Manly, J. J. (2015). The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 70(4), 557-567.
- Stålhammar, J., Hellström, P., Eckerström, C., & Wallin, A. (2022). Neuropsychological test performance among native and non-native Swedes: Second language effects. *Archives of Clinical Neuropsychology*, 37(4), 826-838.
- Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology*, 11(11), 1006-1012.
- Warrington, E. K. (1991). Visual object and space perception battery. (*No Title*).
- Wasserman, J. D. (2017). Nonverbal neuropsychological assessment. In *Handbook of nonverbal assessment* (pp. 287-310). Cham: Springer International Publishing.
- Wechsler, D. (1945). A Standardized Memory Scale for Clinical Use. *The Journal of Psychology*, 19(1), 87-95. DOI: <https://doi.org/10.1080/00223980.1945.9917223>
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale - Fourth Edition*. Bloomington, MN: PsychCorp, an imprint of Pearson Clinical Assessment.
- Yaffe, K., Falvey, C., Harris, T. B., Newman, A., Satterfield, S., Koster, A., ... & Simonsick, E. (2013). Effect of socioeconomic disparities on incidence of dementia among biracial older adults: prospective study. *Bmj*, 347.

- Zahodne, L. B., Manly, J. J., Smith, J., Seeman, T., & Lachman, M. E. (2017). Socioeconomic, health, and psychosocial mediators of racial disparities in cognition in early, middle, and late adulthood. *Psychology and aging, 32*(2), 118.
- Zahodne, L. B., Sol, K., & Kraal, Z. (2019). Psychosocial pathways to racial/ethnic inequalities in late-life memory trajectories. *The Journals of Gerontology: Series B, 74*(3), 409-418.
- Zahodne, L. B., Sharifian, N., Kraal, A. Z., Zaheed, A. B., Sol, K., Morris, E. P., ... & Brickman, A. M. (2021). Socioeconomic and psychosocial mechanisms underlying racial/ethnic disparities in cognition among older adults. *Neuropsychology, 35*(3), 265.

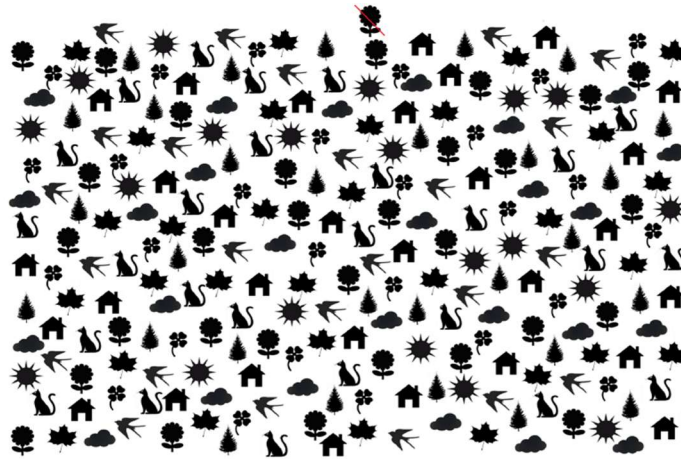
## APPENDIX

Protocol and stimuli of the Non-Verbal Assessment (NoVA) Battery Test.

<p><b>Non-Verbal Assessment Battery -NoVA: batteria non verbale per la valutazione cognitiva</b>          D'Ippolito A., Vestri A., Pucci V., Colombo I., Sebastianutto G., Mondini S.,</p>	
Cognome e Nome _____ Data di Nascita _____ Et� _____	
Data Valutazione _____ Sesso M <input type="checkbox"/> F <input type="checkbox"/> Scolarit� _____ Professione _____	
Manualit� DX – SX sCRI _____ Motivo _____	

<b>Attenzione Visuo-Spaziale</b>	
Barrage di Fiori	/30
TMT-A	
TMT-B	
Symbol Figure Modalities Test	/54
<b>Abilit� Visuo-Costruttive</b>	
Copia di disegno bidimensionale	/2
Copia di disegno tridimensionale	/2
<b>Programmazione Motoria</b>	
Imitazione gesti	/3
Imitazione sequenze motorie	/3
<b>Agnosia Visiva</b>	
Riconoscimento di Figure Incomplete	/4
<b>Memoria</b>	
Color Span Visivo – Avanti	/8
Color Span Visivo – Indietro	/8
Memoria Visiva Immediata	/10
Memoria Visiva Differita	/10
<b>Ragionamento</b>	
Relazioni logiche	/4
Interferenza	/4
Go/No-Go	/4
<b>Punteggio Globale</b>	
-Grezzo	
-Corretto	

Scoring Sheet



Flower Barrage sheet

**Memoria Visiva**

Quello che dovrà fare è memorizzare l'immagine che le mostrerò. Dopodiché la toglierò e le mostrerò altre immagini e deve trovare quella che ha già visto. Vede quest'immagine (mostra primo item)? Bene, la memorizzi (lascia memorizzare per 3 secondi). Bene (mostra rumore per 3 secondi), adesso indichi quali tra queste (mostra target e distrattori) è quella che appena visto.

Memoria Immediata		Memoria Differita		Note
Item	Scoring	Item	Scoring	
	0 1		0 1	
	0 1		0 1	
	0 1		0 1	
	0 1		0 1	
	0 1		0 1	
	0 1		0 1	
	0 1		0 1	
	0 1		0 1	
<b>Totale</b>	<b>/10</b>			

Immagine che contiene schizzo, linea, design, origami

Descrizione generata automaticamente

**Gesti e Sequenze Motorie**

**Imitazione gesti**

Adesso le mostrerò dei gesti con la mia mano che lei dovrà copiare. Per esempio, se io faccio così (riprodurre il primo item con la mano), lei cosa farà? Dopo essersi accertato che l'esaminatore ha compreso, l'esaminatore dirà: bene, continuiamo. Se faccio questo (secondo item) lei farà? E così via

Item	Totale	Scoring
	/3	0 1
<b>Totale</b>		

**Imitazione sequenze motorie**

Adesso le mostrerò una serie di movimenti che deve osservare. Dopo averli osservati eseguirà i movimenti con me e poi dovrà continuare da solo/a. Mi raccomando, osservi per bene il movimento e inizi a farlo solo quando le faccio un cenno. Per esempio, osservi quello che faccio (pugno, taglio, piatto\*3). Bene, adesso facciamo insieme (ripetere 3 volte). Bene, adesso continui da solo. NB: Ripetere il movimento tre volte in fase di osservazione, tre volte in esecuzione congiunta e far eseguire sei volte il movimento in autonomia.

Item	Scoring
Pugno - Taglio - Piatto	0 1 2 3
<b>Totale</b>	<b>/3</b>

**Ragionamento**

**Relazioni Logiche**

Qui ci sono delle immagini in cui manca un elemento. In questo compito deve individuare quale tra le quattro immagini qui sotto è quella che completa la griglia. Per scegliere la giusta immagine dovrà individuare la giusta regola di associazione.

Item	Scoring
1	Dimostrativo
2	0 1
3	0 1
4	0 1
5	0 1
<b>Totale</b>	<b>/4</b>

**Controllo esecutivo**

**Interferenza**

Quando vede un pallino nero (indica) batte una volta con la mano (far battere). Quando vede un pallino rosso (indica) batte due volte con la mano (far battere). Bene, uniamo le due cose: quando vede un pallino nero batte una volta, quando vede un pallino rosso batte due volte. Iniziamo:

Scoring	
0	fino a 3 corrette
1	fino a 7 corrette
2	fino a 11 corrette
3	fino a 15 corrette
4	16 corrette
<b>Totale</b>	<b>/4</b>

**Go/No-Go**

Quando vede un pallino nero (indica) batte due volte con la mano (far battere). Quando vede un pallino rosso (indica) non deve battere. Bene, uniamo le due cose: quando vede un pallino nero batte due volte, quando vede un pallino rosso non deve battere. Iniziamo:

Scoring	
0	fino a 3 corrette
1	fino a 7 corrette
2	fino a 11 corrette
3	fino a 15 corrette
4	16 corrette
<b>Totale</b>	<b>/4</b>

**Agnosia Visiva**

**Riconoscimento di Figure Incomplete**

Qui vede una figura in parte cancellata ma ancora riconoscibile. Dovrà trovare tra le immagini quella uguale a quella con il contorno cancellato.


Item	Scoring
1	Dimostrativo
2	0 1
3	0 1
4	0 1
5	0 1
<b>Totale</b>	<b>/4</b>

**Copia di Disegno**

Item	Scoring
Dis. Bidimensionale	0 1 2
Dis. Tridimensionale	0 1 2

Administration protocol page, including instructions and scoring sections for the different tasks of the NoVA Battery

### Color Span Visivo – Avanti

Adesso le mostrerò dei colori in un ordine specifico. Dopo che ha visto la sequenza, le mostrerò un'immagine con dei colori. Quello che dovrà fare è toccare i colori nello stesso ordine in cui li ha visti. Inizieremo con pochi colori e aumenteremo man mano. Facciamo una prova. Se le mostro questi (avvia animazione), lei toccherà? Esempio: 

	3		6
	3		6
	4		7
	4		7
	5		8
	5		8

### Color Span Visivo – Indietro

Adesso le mostrerò dei colori in un certo ordine. Dopo che ha visto la sequenza, le mostrerò un'immagine con dei colori. Quello che dovrà fare è toccare i colori NELL'ORDINE INVERSO, dall'ultimo al primo, rispetto a quello che le ho mostrato. Inizieremo con pochi colori e aumenteremo man mano. Facciamo una prova e si ricordi che dobbiamo partire dall'ultimo colore visto andando indietro. Se le mostro questi (far partire animazione), qual è l'ordine inverso?

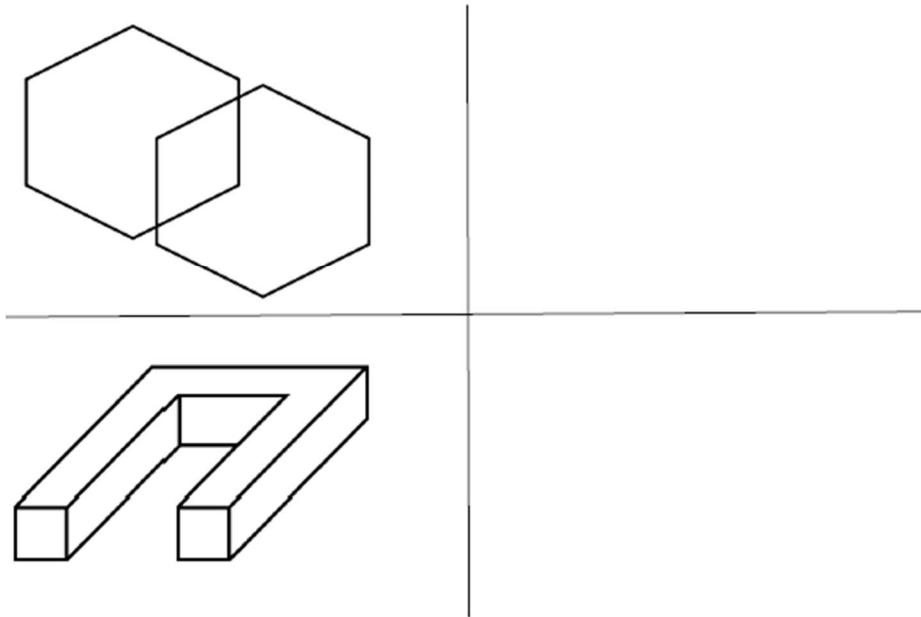
Esempio:  -> 

	3		6
	3		6
	4		7
	4		7
	5		8
	5		8

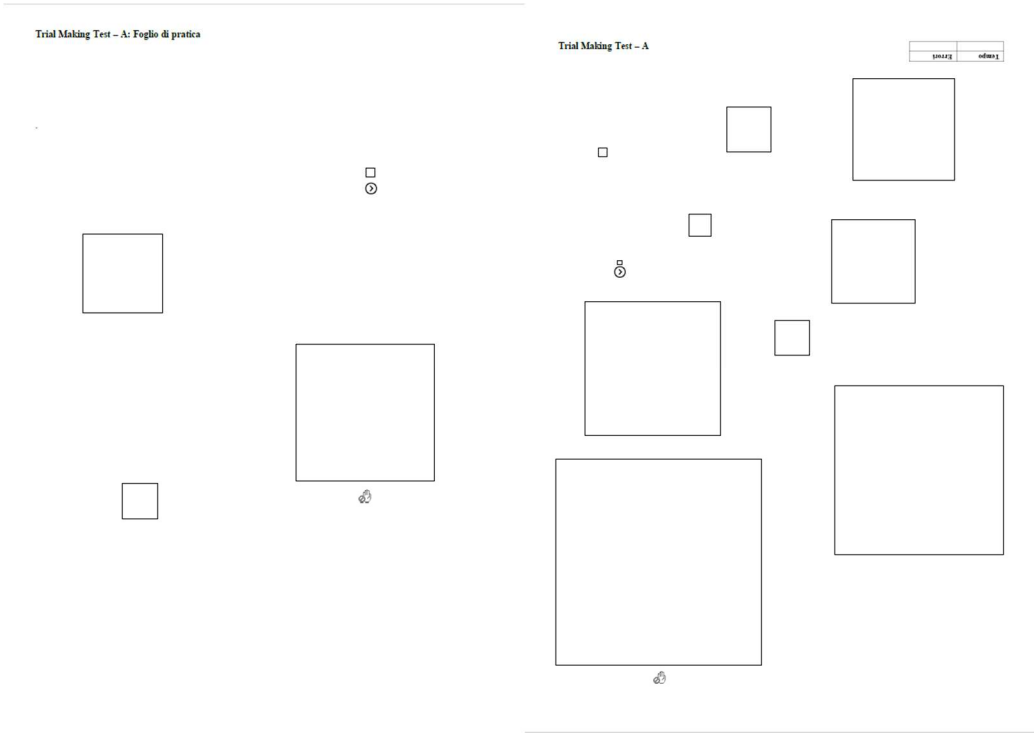
### Abilità Visuo-Costruttive

#### Copia di Disegno

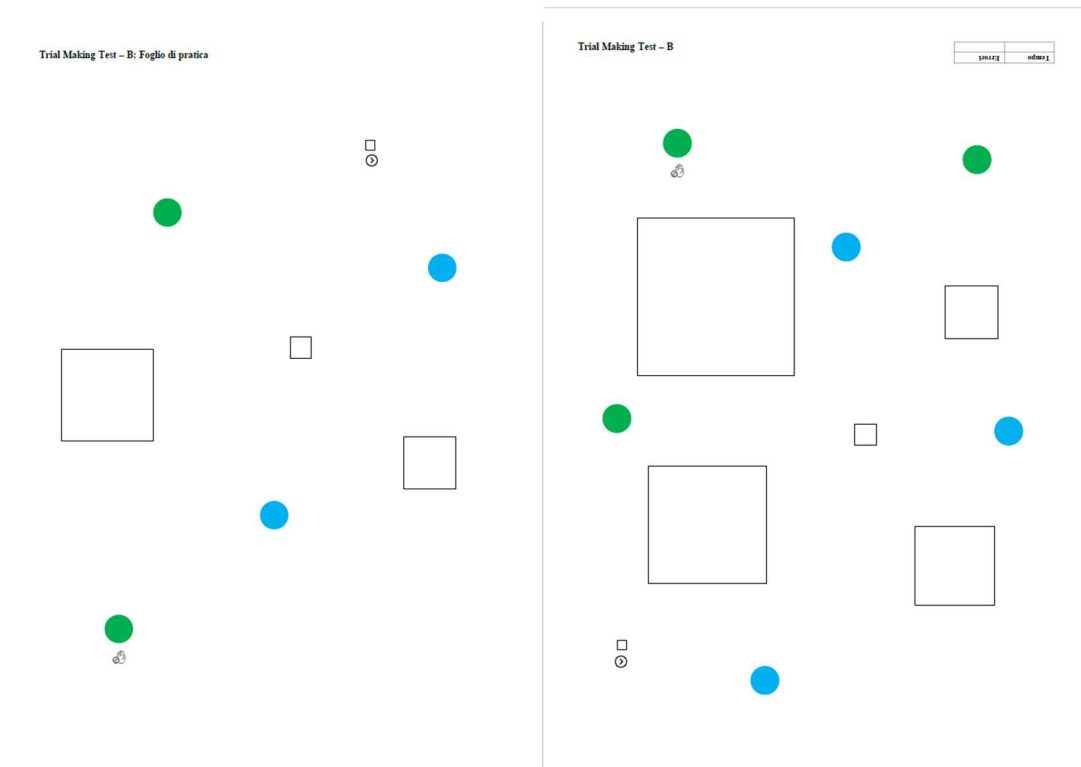
Vede queste immagini (indica)? Provi a ricopiarle qui a lato (indica) il più uguale possibile.



Page including instructions and scoring sections for the Visual Colour Span task and the Drawing Copying tasks





















TMT-A's practice sheet and test sheet


































































TMT-B's practice sheet and test sheet

**Symbol Figure Modalities Test**

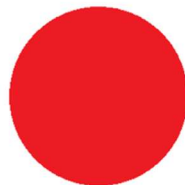
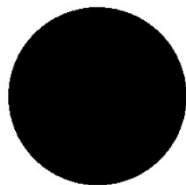
*Qui ci sono varie figure (indica le silhouette) a cui sono associati dei simboli (indica i simboli). Come vede alla prima figura (indica) è associato questo simbolo (indica), alla seconda (indica) questo simbolo (indica) e così via. Quello che deve fare è scrivere negli spazi vuoti (indica) il simbolo corrispondente alla figura. Per esempio, per la prima figura (indica rondine) dovremmo scrivere questo simbolo (indica X e scrive), per la casa (indica) dovremmo scrivere questo simbolo (indica O e scrive). Adesso provi lei. Attenzione perché avrà a disposizione 90 secondi, quindi deve cercare di essere il più veloce possibile. Iniziamo*

---

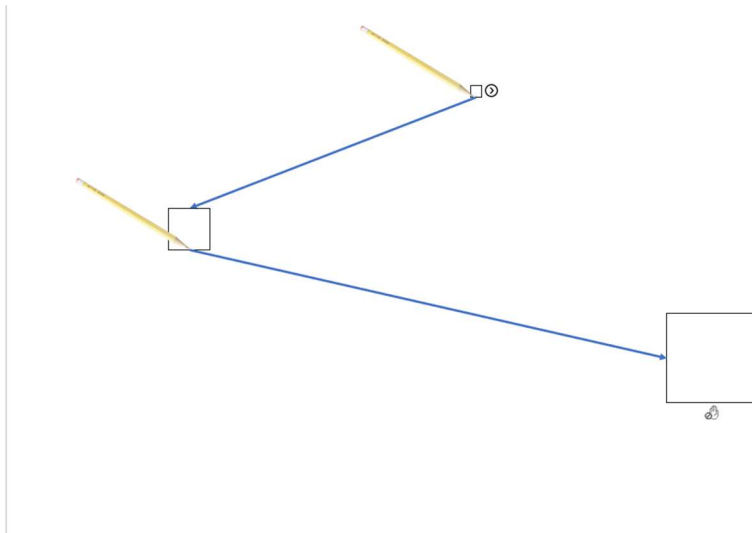
Symbol Figure Modalities Test administration sheet with instructions



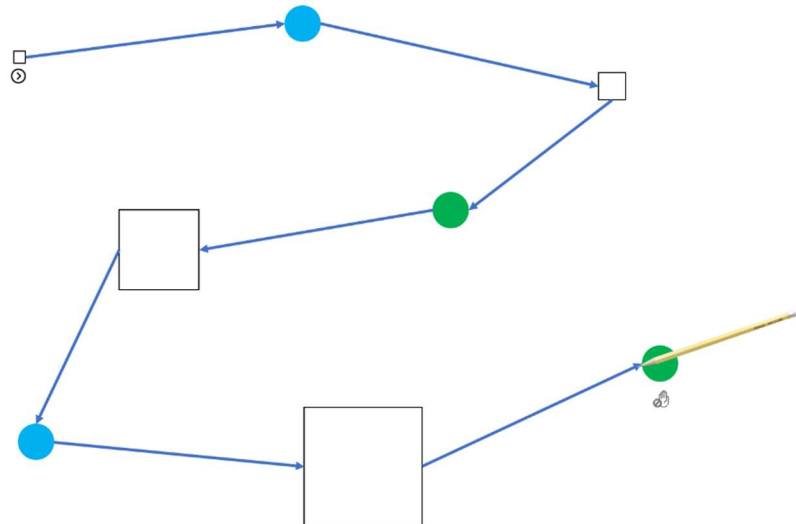
Stimuli for Interference and Go/No-Go tasks



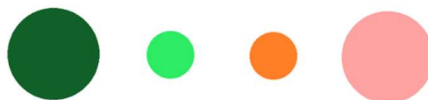
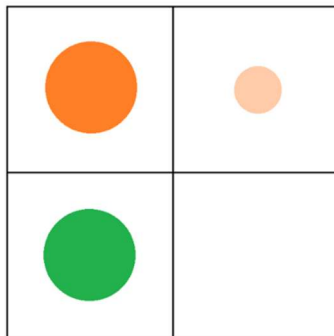
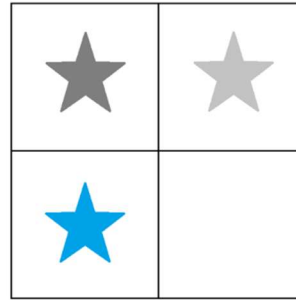
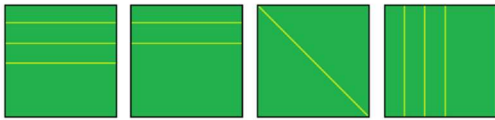
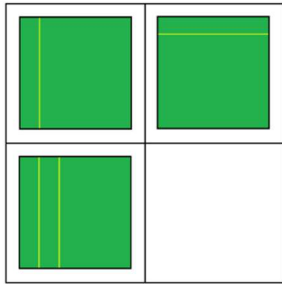
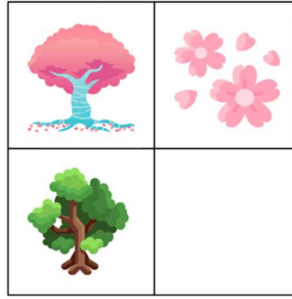
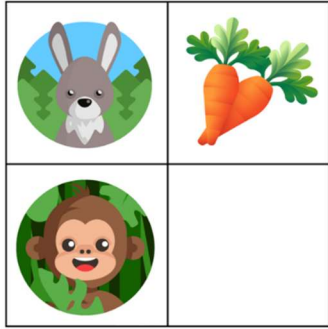
Screen display of the Visual Colour Span task



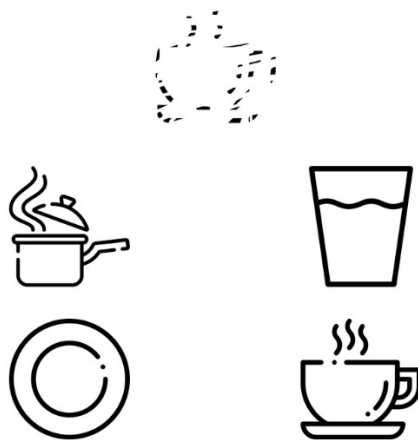
TMT-A demonstration slide



TMT-B demonstration slide



Example slide and test slide for the Logical Relations task



Example slide and test slide for the Incomplete Figures task