

# Italian executive functions battery for preschoolers (BAFE): working memory, inhibition, set-shifting

Paolo Stievano<sup>1</sup> - Matteo Ciancaleoni<sup>2</sup> - Giovanni Valeri<sup>3</sup>

1. Azienda Sanitaria Locale, Roma 4, Rome, Italy

2. Hogrefe Editore, Florence, Italy

3. Department of Neurosciences and Neurorehabilitation, IRCCS Children Hospital Bambino Gesù, Rome, Italy

DOI: <http://dx.doi.org/10.7358/neur-2017-022-stie>

[paolo.stievano@inwind.it](mailto:paolo.stievano@inwind.it)

---

## ABSTRACT

*This study analyzes the psychometric characteristics of an Italian test of executive function (EF) tests assessing inhibition, set-shifting and working memory (core executive functions) for preschool children. BAFE is based on a concept accredited by the majority of EF research: a representation articulated in a profile of neuropsychological processes relatively independent. The main goal of this study was to assess the multiple EF components of BAFE Test in 3 to 6-year-old children and its sensitivity, validity and reliability to age differences. In order to evaluate the convergent validity of the BAFE test, a correlation between different tests assessing similar neuropsychological constructs were performed. The framework of the relationships between the tests for children shows associations and dissociations compatible with the theoretical hypotheses on executive functions. BAFE test offers additional measures for clinical evaluations for preschoolers. It could introduce new clinical practice that add to clinical observations or questionnaires a rigorous measurement tools with good psychometric properties.*

*Keywords: Executive functions; Inhibition; Working memory; Set-shifting; Developmental neuropsychology*

## 1. INTRODUCTION

The aim of this study is to provide an Italian validation of executive function tests assessing inhibition, set-shifting and working memory for preschool children.

The tests we propose, BAFE (Valeri, Stievano, Ferretti, Mariani, & Pieretti, 2015), can be easily administered to preschool age children due to its brief, simple, and enjoyable layout, which resonates closely to the interests of the children.

Executive control is a term used to refer broadly to those cognitive abilities that are associated with or sub served by the prefrontal cortex and the interconnected subcortical system (Diamond, 2013; Stuss, 1992).

Defined as the regulation of more automatic processes (thoughts, behaviour, emotion) in the service of a goal the construct of executive function has gained prominence in the developmental literature in the past few decades (Zelazo & Muller, 2002).

This topic is very useful to understand typical and atypical development of children. The implications are important in clinical and educational field.

Prefrontal systems undergo a protracted course of development (Benes, 2001). In comparison with posterior cortical areas, the phases of prefrontal cortical development, including neuronal generation, differentiation and synaptic pruning, occur later and over a longer period of time (Shaw et al., 2006).

Executive functions (EF) are critically important in the overall neuropsychological functioning of the developing child and play a crucial role in the child's cognitive, behavioural, and social-emotional development. Early EF ability has been associated with various measures of subsequent adjustment, including social understanding and peer relations (Hughes & Ensor, 2011); conscience development (Kochanska & Knaack, 2003), and even financial success decades later (Moffitt et al., 2011). Thus, sensitive measures of early EF may provide powerful tools for the early detection and prediction of adjustment. Deficits in executive functions are central characteristics of many acquired and developmental disorders. (Barkley, 1997; Isquith, Gioia, & Espy, 2004; Pennington & Ozonoff, 1996).

A current view of preschooler's behaviour is that children are "Dysexecutive": they are not able to exert higher order control of pertinent cognitive processes, emotional responses, and behavioural impulses, as lack of inhibitory control, significant attentional lability, less cognitive flexibility and lack of organized or planful strategic behaviour and self-monitoring are present in preschool age.

The earliest roots of developmental disorders, evident in later childhood and adolescence, can be found by studying the emergent executive functions in preschoolers. For example, better understanding of the roots of poor inhibitory control, later manifested in attention-deficit/hyperactivity disorder (Sonuga-Barke, 2003).

A central issue in the assessment of early EF concerns the form or structure of EF. While some have argued for a unitary construct with subprocesses (Baddeley, 2000; Baddeley & Wilson, 2002), others have argued for multiple components (Pennington & Ozonoff, 1996). Miyake, Friedman, Emerson, Witzki, and Howerter (2000) proposed a potential solution by integrating the two perspectives. Using confirmatory factor analysis (CFA), they found evidence for three moderately correlated EF factors (working memory, inhibition and shifting) in adults.

The findings for pre-schoolers have been mixed with some research supporting a unitary EF factor (Wiebe et al., 2011) and some supporting a two factor structure with *inhibition* and *working memory* as separate correlated factors (Mueller, Kerns, & Konkin, 2012). Upon reviewing the studies on EF development from infancy to 5 years, Garon, Bryson, and Smith (2008) extended Miyake et al. (2000) integrative EF framework by proposing that EF develops in a hierarchical manner with attention serving as the foundation. According to Garon et al. (2008) key EF components, emerge during the first 3 years of life from early simple skills such as holding information in mind and delaying in a response. Further, these simpler components become integrated into the complex processes (i.e. inhibition, working memory, and shifting) that characterize mature EF abilities. These authors hypothesized that more complex EF abilities that develop later, such as shifting and planning, are constructed from earlier developing EF abilities. From core EFs (working memory, inhibition, flexibility), higher-order EFs are built such as reasoning, problem solving, and planning (Collins & Koechlin, 2012; Lunt et al., 2012).

## *1.1 Core executive functions*

### *1.1.1 Set shifting*

Set shifting or mental flexibility refers to the ability to shift between different thoughts and actions according to changes in a situation, utilize feedback, generate concepts, and devise alternative problem-solving strategies (Lezak, 2004). Brain regions that have been linked to these functions include the lateral prefrontal, orbitofrontal, and parietal cortices, basal ganglia, and cerebellum (Howard, Anderson, & Taylor, 2010). Children with problems in mental flexibility appear inflexible or rigid and they have difficulties moving between activities and adjusting to new situations or change routine.

Toward the end of the preschool period, however, children become able to reflect on rules, integrating conflicting elements of knowledge into a more complex rule system (Zelazo, Qu, Muller, Schneider, Schumann-Hengsteler, & Sodian, 2005)

Diamond emphasizes that difficulty overcoming conflict is the root cause

of perseverative behaviour in young children. She describes EF as the ability to overcome automatic, prepotent behaviour despite the pull of previous experience (Diamond, 2014)

Set-shifting tasks involve shifting from one “mental set” to another. Regardless of the particular form that they take, all set-shifting tasks involve two phases. The first phase requires participants to form a mental set in which an association is made between a particular stimulus and a response. In forming this set, participants must focus on relevant stimuli and ignore distractors and then hold the mental set (rule) in working memory. Tasks differ on this initial working memory load. The second phase of these tasks involves shifting to a new mental set that in some way conflicts with the first. Tasks may therefore also differ in the amount of conflict that participants have to overcome. Probably the most relevant distinction that has been made in the literature concerns the nature or type of shift required. Shift type is determined by whether the conflict occurs at the perceptual or response stage.

The ability to resolve conflicts during information processing is hypothesized to be one of the results of this maturing anterior attention system (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005). This ability has been taken as the most important milestone in the development of EFs (Rueda et al., 2005). Conflict tasks, in fact, have been regarded as a good measure of how attention directly influences EF (Rueda et al., 2005). Moreover, in most models of cognitive control, the ability to resolve conflicts is essential (Cohen, Aston-Jones, & Gilzenrat, 2004). Although the mechanism underlying this is still not clear, most theorists have suggested that attention networks resolve conflicts by modifying the activation of other brain networks (Miller & Cohen, 2001). As such, the ability to resolve conflict may be seen as a special function of selective attention, which helps children increase focus on a particular stimulus in the service of task demands. The ability to resolve conflicts matures slowly in the first 2 years, showing noticeable decreases in development after this period until about 6 years of age (Rueda et al., 2005).

### *1.1.2 Inhibition*

The term “Inhibition” or “Inhibitory control” encompasses the ability to suppress a prepotent response, interrupt an ongoing response, and resist distraction from external stimuli (Barkley, 1997). These abilities are thought to be associated with a distributed neural network that includes the lateral prefrontal cortex, anterior cingulate, and basal ganglia (Tamm, Menon, & Reiss, 2002). Choice would not be possible were we not able to resist, at least partially, the pull of external stimuli, our emotions, or old habits of mind or behaviour. Inhibitory control involves being able to control one’s attention,

behaviour, thoughts, and/or emotions to override a strong internal predisposition or external lure. Inhibitory control of attention (selective or focused attention) enables us to focus on what we choose, suppressing attention to other stimuli. Discipline is the aspect of inhibitory control that involves making yourself do something, or keep at something, though you would much rather be doing something else (Diamond, 2014).

### *1.1.3 Working memory*

Working memory, the capacity to temporarily hold and manipulate information for the purpose of guiding future responses and behaviours (Baddeley, 2000), is subserved by an integrated neural network that includes the dorsolateral prefrontal cortex, anterior cingulate, parietal and temporal cortices, hippocampus, and basal ganglia (Luciana & Nelson, 1998).

Reasoning involves holding information in mind and seeing how they relate. Working memory allows to hold information present in mind and mentally working with it; and permit to manage information no longer perceptually present too. Working memory is critical for making sense of anything that unfolds over time, for that always involves relating what came earlier to what came later. Working memory is critical to our ability to see connections between seemingly unrelated things. Working memory enable us to bring conceptual knowledge and not just what is perceptually present to bear on our decisions, and consider the past and our future hope in making plans and decisions (Diamond, 2014).

Children with deficits in working memory often have difficulty remembering things even for a few seconds and tend to lose track of what they are doing as they work.

Children in preschool age at risk of ADHD, preterm, language impairment, epilepsy, traumatic brain impairment often show EF deficits.

The latest studies on the development of executive functions are interesting for understanding the typical and atypical development of the child. The EF role is crucial to improve the effectiveness of clinical and rehabilitative interventions in children.

However, clinical application of experimental evidence is still limited. Often, the assessment of executive functions is performed only by clinical observations or disclosure of information through questionnaires completed by teachers and/or caregivers.

The BAFE battery project aims to promote clinical approaches supported by psychometrically valid test is also, supplied the opportunity to guide the rehabilitation projects towards programs aimed at increasing lacking specific processes.

For these reason, we decided to elaborate and standardize a battery of

three core areas of EF (inhibition, set-shifting, working memory) in order to offer additional measures for clinical evaluations for pre-schoolers.

The overarching goal of this study was to assess the multiple EF components of our EF battery in 3 to 6-year-old children and its sensitivity, validity and reliability to age differences. Furthermore, we were interested in exploring the structure of EF components in preschoolers using this battery.

In order to evaluate the convergent validity of the BAFE test, a correlation between different tests assessing similar neuropsychological constructs were performed. We predicted a relationship between the measures of different tests measuring similar construct (e.g. inhibition, set-shifting, and working memory).

We also wanted to study divergent validity to ensure the specificity of the test battery BAFE and its evaluation of neuropsychological processes functionally independent from subdomains of EF, language (Peabody test) and cognition (Leiter-R). For this reason, we expected no correlations between these measures.

## 2. METHODS

### 2.1 Participants

Participants were 358 Italian children (53% male, age range from 3 to 6 years,  $M = 53.56$ ,  $SD = 9.92$  months). Due to the nature of the scale, Pattern Making scale was not administered to children of 5 years of age or above. This task was proposed only to children of the first three bands of age. In fact, after five years there is a ceiling effect for attentional flexibility.

Reliability and item analysis were calculated using a subgroup of 101 children due to the lack of data on the items for all children. The subgroup is composed by children (Male = 53%) from 3 to 6 years with a mean age of 58.19 months ( $SD = 9.99$  months); therefore, this subgroup is highly comparable with the total sample for age and gender.

### 2.2 Material

These specific psychometric measures or tests have been selected according to a functional perspective of cognitive domains depending on different areas of the prefrontal cortex. The EF tasks in particular are linked to specific subdomains according to the unitary construct of EF with partially dissociable components (Garon, Bryson, & Smith, 2008).

### *2.2.1 Inhibition tasks*

DAY AND NIGHT “Stroop-like day-night task” (Carlson, 2005; Gerstadt, Hong, & Diamond, 1994; for the Italian adaptation, see Valeri et al., 2015). The examiner engaged the children in conversation about when the sun comes up (in the day) and when the moon and stars come out (in the night). He then presented a white card with a drawing of a yellow sun and a black card with drawing of a white moon and stars. Children were then told that the instructions for this game required them to say “night” for the sun card and “day” for the moon/stars card. Accuracy (number correct out of 16) was recorded.

### *2.2.2 Set-shifting task*

PATTERN MAKING TEST (Attentional Flexibility, Valeri et al., 2015). This test is a magnets pattern-making task adapted (Frith, 1971; Hughes, 1998), and used as a measure of frontal function (Hughes, 1998; Passler, Isaac, & Hynd, 1985). Children were first shown a sequence of 18 coloured circles on a long strip of card, and asked to name each colour in turn. The examiner says “Yes, you see it makes a pattern: blue-blue-red, blue-blue-red, emphasizing the words in a rhythmic way. Then the examiner instructs the child to try and make exactly the same pattern on a steel rule using a set of red and blue magnets. No feedback is given during the task. Scoring is number of rights triplets (six blue-blue-red). This task is related to a Shift type in which conflict occurs at the perceptual not response stage.

CARD SORT (Carlson & Moses, 2001; Frye, Zelazo, & Palfai, 1995; Zelazo, 2004; for the Italian adaptation see Valeri et al., 2015). Children were introduced to two recipe boxes that had rectangular slots cut in the top. Target cards (e. g. red bear, blue house) were affixed to the front of the boxes. The examiner presents a series of cards (red and blue bear and houses) and instructed children to place all the bears in the box with the red bear and to place all the houses in the box with the blue house in the “shape game”. After five consecutively correct trials, the experimenter announced that they would stop playing the shape game and now play the “color game”. In this case, all the red things would go in the box with the red bear affixed and all blue things would go in the box with the blue house affixed. The total number of correct incompatible postswitch trials was recorded. This task is related to a Shift type in which conflict occurs at response stage.

### 2.2.3 Working memory task

SPIN THE POTS (Diamond & Taylor, 1996; Hughes, 1998; for the Italian adaptation see Valeri et al., 2015). In this task the child was asked to place an object (a red ring) in each of the eight different pots arranged on a tray. The tray was then covered with a scarf and spun around. The child was then asked to lift the scarf and choose a pot. Each time the child chose a baited pot, the sweet or the object was put into a small reward envelope for the child to keep. This procedure was repeated until eight smarties had been found or after 15 trials had been conducted (whichever was sooner). Performance was rated by the number of trials required to find all eight objects.

### 2.2.4 Inhibition test (*Tests measuring convergent validity of the battery*)

WHISPER (Carlson, 2005; Kochanska, Murray, Jacques, Koenig, & Vandegest 1996). The examiner asked children if they could whisper their names and then presented a series of 10 cards depicting cartoon characters (6 familiar, 4 unfamiliar to most preschoolers). Children were told to whisper their names of each character and that it was okay if they did not know all of them. On each trial they received a score of 0 if they blurted out the name or used a normal voice and a score of 1 if they whispered. Unfamiliar characters were included so that children would be more excited upon seeing a familiar one (and more likely to shout out the name); “don’t know trials were unscored.

NEPSY (Korkman, Kirk, & Kemp, 2007; for the Italian Adaptation see Urgesi, Campanella, & Fabbro, 2011) NEPSY subtest Statue was used where the child was asked to stand still in silence with the eyes closed posing as a “statue” while pretending to hold a flag for over a timeframe of 75 seconds.

BEAR/DRAGON (Carlson, 2005; Reed, Pien, & Rothbart, 1984) The examiner introduced children to a “nice” bear puppet (using a soft, high-pitched voice) and a “naughty” dragon puppet (using a gruf, low-pitched voice). He explained that in this game they are to do what the bear asks them to do (e.g. touch your nose) but not to do what the dragon asks. After practicing, there were 10 test trials with the bear and dragon commands in alternating order. Children were seated at a table throughout the task, and all actions involved hand movements. Performance on dragon trials was taken as an index of self control (0 = movement, 1 = no movement, scored individually for each trial).

### 2.2.5 Verbal ability measure

Children were given the Peabody Picture Vocabulary Test, revised edition

(PPVT-R) (Dunn & Dunn, 1981). The PPVT-R measures an individual's receptive (hearing) vocabulary. It provides a quick estimate of verbal ability.

The PPVT- R is a measure of receptive vocabulary that correlates highly with full-scale verbal intelligence measures such as the Wechsler Preschool and Primary Scales of Intelligence (revised, WPPSI-R) (Carlson, Moses, & Claxton, 2004) and the verbal subscale of the Stanford–Binet IV (Carlson et al., 2004; Hodapp, 1993), as well as with theory of mind (Carlson & Moses, 2001, Carlson et al., 2004).

### *2.2.6 Cognitive abilities*

The Leiter International Performance Scale-Revised (Leiter-R) (Roid & Miller, 1997) is a standardized nonverbal measure of intelligence. The Leiter-R evaluates nonverbal cognitive, attentional and neuropsychological abilities and targets “typical” as well as “atypical” children, adolescents, and adults.

We used the “Brief IQ”, based on four subtests (Repeated Patterns, Sequential Order, Figure–Ground, and Form Completion).

### *2.2.7 Planning measures*

The Tower of London (TOL) was originally developed by (Shallice, 1982) in an effort to assess executive planning.

## 3. PROCEDURE

Each participant individually completed the battery in a standardized format. The administration time ranged from 10 to 20 minutes with an average time of 16 minutes. Children were tested in a small quiet area of their school by two trained experimenters (one psychologist and one speech therapist) graduate. The scoring was double-checked by the psychologist who administered the tests and by the research team leader (the first author of the article). Participants' parents signed an informed consent document that assured them that the data would be handled to protect the participants' privacy. The study protocols were approved by the institutional review boards at each school.

Data analysis was performed to check item quality, reliability and validity of the instrument.

#### 4. ITEM ANALYSIS

In order to evaluate items, two different types of analyses were conducted; Item difficulty and Item discrimination. In testing item difficulty, the correct answer percentage was calculated with values between .20 and .80 regarded as adequate (Crocker & Algina, 1986). For Item discrimination, power was tested using D index and a corrected item-total correlation. D index is the difference in item difficulty between high ability subgroup, i.e. people who scored within the 70° percentile and above, and low ability subgroup, i.e. people who scored within the 30° percentile and below. D index values of .20 or above are acceptable (Wiersma & Jurs, 1990). Due to nature of the scale, item analysis was not performed for Spin the Pots scale.

##### *4.1 Dimensionality*

The dimensionality of the test was tested with a Confirmative Factor Analysis (CFA) for dichotomous data using Mplus 3.0 software (Muthén & Muthén, 2004) that implemented the Weighted Least Squares Means and Variance adjusted (WLSMV) estimation method. WLSMV uses weighted least square parameter estimates from the diagonal of the weight matrix. This method is recommended for categorical variables (Muthén, 2001; Muthén & Muthén, 2004; Flora & Curran, 2004) on the basis of simulation studies (Muthén, duToit, & Spisic, 1997). In order to evaluate the fit of the model, the chi square, the Comparative Fit Index (CFI) (Bentler, 1990), the Tucker-Lewis Index (TLI) (Tucker & Lewis, 1973) and the Root Mean Square Error of Approximation (RMSEA) (Browne & Cudeck, 1983) statistics were used. Bentler (1990) and Bentler and Bonnet (1980) suggested that values higher than .90 for the CFI and TLI indicate that the model provides an adequate fit to the data and Browne and Cudeck (1983) suggested that values of RMSEA lower than .05 indicate a close fit. Due to nature of the scale, dimensionality was not tested for Spin the Pots scale.

##### *4.2 Reliability and validity*

The reliability of the battery was tested using KR-20 index due to the dichotomous format of the data.

Validity of the test was tested in different ways. First of all, in order to evaluate the fairness of the test differences between administrators were assessed. Moreover, to clarify that the test did not advantage male or female, differences across gender were examined. Due to the relation of the construct with the age, the differences between age subgroups were tested. In order to do so, a univariate ANOVA was performed with Scheffé post-hoc analysis and a  $\eta^2$  effect size index.

## 5. RESULTS

### *5.1 Item analysis*

Item analysis attested the good item quality. Item varied from .63 to .93 for difficulty. However, the results also show that there are more easy items than difficulty ones. Moreover, items showed good discrimination power such that all items had a D index of .20 or above and corrected item-total correlations ranged from .29 to .86.

As a first step, we tested the dimensionality of the battery in order to evaluate the factorial structure of the test. In line with the theoretical framework which guided the development of the BAFE, a three factor structure was tested. Specifically, the CFI = .99 and the TLI = .98 indicated a very good fit (Bentler, 1990) as well as the RMSEA = .06 (90% CI: .03 - .09) (Browne & Cudeck, 1993). Factor loadings were all significant ( $p < .001$ ).

### *5.2 Reliability*

KR-20 indices showed that the three scales had an appropriate reliability. In detail, KR-20 indices were higher than the cutoff value of .70 (KR-20 = .77, .92 and .92 for Card Sort, Night & Day, and Pattern Making, respectively). Therefore, the Night & Day and Pattern Making measures particularly showed high internal coherence.

### *5.3 Validity*

In order to assess the test fairness, differences between administrators were computed. Due to the fact that only two researchers administrated the majority of the tests, comparisons were performed solely between the two administrators (labelled 1 and 2). Results showed no differences in each scale (Card Sort:  $t_{(299)} = -.43$ ,  $p > .05$ ; Night & Day:  $t_{(296)} = -1.94$ ,  $p > .05$ ; Pattern Making:  $t_{(120.14)} = 1.92$ ,  $p > .05$ ; Spin the Pots:  $t_{(296)} = .75$ ,  $p > .05$ ) demonstrating the fairness of the test.

To evaluate the presence of gender differences across the four scales of the battery, t test analyses were performed. Results showed no differences for each scale (Card Sort:  $t_{(356)} = -.15$ ,  $p > .05$ ; Night & Day:  $t_{(353)} = -.58$ ,  $p > .05$ ; Pattern Making:  $t_{(209)} = 1.48$ ,  $p > .05$ ; Spin the Pots:  $t_{(346)} = -.99$ ,  $p > .05$ ) indicating gender equity of the test.

Age differences were tested using univariate ANOVA. Different age subgroups were created such that each subgroup contained children with an age range of 6 months according to literature (Carlson, 2005). Results showed that all differences were significant (Card Sort:  $F_{(5,352)} = 7.37$ ,  $p < .001$ ,  $\eta^2 = .094$ ; Night & Day:  $F_{(5,349)} =$

10.88,  $p < .001$ ,  $\eta^2 = .134$ ; Pattern Making:  $F_{(3;207)} = 29.29$ ,  $p < .001$ ,  $\eta^2 = .298$ ; Spin the Pots:  $F_{(5;342)} = 6.35$ ,  $p < .001$ ,  $\eta^2 = .085$ ). In line with Cohen's (1988) criteria, the eta square indices showed that the differences ranged from average to elevate. Finally, post hoc analyses were performed using Scheffé method. Results showed that differences between younger and older children were significant.

Table 1. Item analysis

Test	Item	Difficulty	D index	Corrected item-total correlation
Card Sort	1	.79	.66	.65
	2	.93	.22	.32
	3	.93	.22	.29
	4	.77	.72	.80
	5	.77	.72	.67
Night & Day	1	.77	.44	.47
	2	.88	.37	.65
	3	.72	.44	.47
	4	.85	.47	.75
	5	.85	.37	.67
	6	.80	.50	.63
	7	.68	.55	.44
	8	.77	.45	.49
	9	.79	.48	.64
	10	.82	.53	.77
	11	.80	.54	.69
	12	.80	.51	.72
	13	.68	.60	.56
	14	.80	.56	.61
	15	.77	.64	.64
	16	.78	.66	.64
Pattern Making	1	.81	.60	.80
	2	.79	.67	.80

*Table 2. Differences between administrators*

<b>Scale</b>	<b>Administrator</b>	<b>N</b>	<b>M</b>	<b>Sd</b>
Card Sort	1	185	2.35	1.04
	2	116	2.40	.90
Night & Day	1	185	12.53	5.01
	2	113	13.61	4.01
Pattern Making	1	117	4.64	1.97
	2	66	4.00	2.26
Spin the Pots	1	185	12.11	3.67
	2	113	11.80	3.11

*Table 3. Gender differences*

<b>Scale</b>	<b>Gender</b>	<b>N</b>	<b>M</b>	<b>Sd</b>
Card Sort	Male	167	2.42	.95
	Female	191	2.43	.95
Night & Day	Male	167	12.99	4.53
	Female	188	13.27	4.32
Pattern Making	Male	97	4.57	1.97
	Female	114	4.14	2.18
Spin the Pots	Male	162	11.57	3.41
	Female	186	11.94	3.38

Table 4. Age group differences

Scale	Age group	N	M	Sd
Card Sort	3.0 – 3.5a,b,c	53	2.08	1.14
	3.6 – 3.11d,e,f	51	1.94	1.19
	4.0 – 4.5	61	2.38	.93
	4.6 – 4.11a,d	60	2.70	.72
	5.0 – 5.5b,e	89	2.62	.75
	5.6 – 6.0c,f	44	2.73	.69
Night & Day	3.0 – 3.5a,b	50	11.30	5.10
	3.6 – 3.11c,d	51	11.39	5.26
	4.0 – 4.5e,f	61	11.72	5.83
	4.6 – 4.11	60	13.43	3.99
	5.0 – 5.5a,c,e	89	14.84	1.80
	5.6 – 6.0b,d,f	44	15.36	.92
Pattern making	3.0 – 3.5	48	2.67	2.72
	3.6 – 3.11	51	3.73	2.03
	4.0 – 4.5	59	5.03	1.61
	4.6 – 4.11	53	5.66	.96
	5.0 – 5.5	---		
	5.6 – 6.0	---		
Spin the Pots	3.0 – 3.5a,b	46	13.07	3.52
	3.6 – 3.11c,d	51	12.84	3.70
	4.0 – 4.5	58	11.69	3.11
	4.6 – 4.11e	60	12.40	3.60
	5.0 – 5.5a,c	89	10.88	3.03
	5.6 – 6.0b,d,e	44	10.20	2.58

Note: a,b,c,d,e,f indicate Scheffé post-hoc differences ( $p < .05$ )

## 6. DISCUSSION

The present data extend previous research on EF in preschoolers by showing that our EF battery (BAFE) is sensitive to cross-sectional age differences from 3 to 6 years. All of EF measures were significantly associated with age.

The framework of the relationships between the tests for children shows associations and dissociations compatible with the theoretical hypotheses set out above. In order to test validity, we investigated the relations between the individual battery BAFE tests and other tests that measure the same neuropsychological constructs or other cognitive skills (e.g. language and non verbal IQ).

The test set-shifting Card Sort had significant correlations with evidence of behavioral inhibition Whisper, but not with other evidence of inhibition contained in the battery Nepsy2. Such dates show that there are probably several forms of behavioral inhibition separable and independent. Furthermore, the process flexibility is saturated even of certain aspects of behavioural inhibition. In line with the theoretical hypotheses, set-shifting ability is independent from a test of cognitive assessment of non-verbal skills (Leiter-R).

Finally, it is interesting to note that in preschoolers, cognitive flexibility and visuospatial working memory are connected with language development (Peabody test). This is most likely due to the mental representation of the child becoming more enhanced and effective through the use of symbolic aspects of language, constituting good support to the cognitive flexibility and visuospatial working memory.

The inhibition test (Night and Day) confirms a good content validity as it is correlated with other evidence of inhibition (Nepsy2). In line with the main theoretical hypotheses on executive function, it appears to be independent of the cognitive skills measured by the test Leiter-R .

Attentional flexibility (Pattern Making Test) is correlated with the test of visuospatial memory (Spin The pots) and planning (Tower of London). Such correlations may have developmental significance when the test planning (Tower of London) is superordinate to other core EF visuospatial working memory and attentional flexibility. Core executive functions (i.e. inhibition, set-shifting, and working memory) are the basis of other EF higher order such as reasoning, problem solving and planning (Collins & Koechlin, 2012; Lunt et al., 2012).

In addition, the collected data are in line with the claims of Teuber (Friedman & Miyake, 2016) about “unity and diversity” to describe frontal lobe functions. Dissociations that we have identified are useful for understanding the structure of EF in the course of their development.

Table 5. Correlations between BAFE scales and other measures

	Card Sort	Night & Day	Pattern Making	Spin the Pot
Whisper	.20*	.19*	.24*	.03
Nepsy2 – Statue – Body	-.12	-.25**	-.32**	.17*
Nepsy2 – Statue – Eyes	-.11	.29***	-.34**	.21**
Nepsy2 – Statue – Voice	-.07	.08	.00	-.06
Nepsy2 – Statue	.14	.28***	.37***	-.20*
Leiter – IQB	.12	-.06	.01	-.16
Peabody	.22*	.16	-.02	-.20*
Bear & Dragon	.15	.25**	.24*	-.09
Tower of London	.18*	.33***	.39**	-.29**

\*\*\*p<.001, \*\*p<.01, \*p<.05

PFC play an important role in coordinating activity across diverse areas. Single case dissociations and low correlations between EF tasks in patient groups support suggestions that executive control might be a collection of processes rather than an entirely unitary function. At the same time, there are some similarities that have led researchers to champion the view that such impairments share something (Friedman & Miyake, 2016).

For example this commonality has been suggested to reflect associations among elements in behavioural inhibition (Barkley, 1997) especially during early child development.

Neuroimaging studies suggest an activation of a unique fronto-cingulo-parietal network during EF tasks (Niendam, Laird, Ray, Dean, Glahn, & Carter, 2012).

The developmental pattern of EF, abilities demonstrated in the present study, is consistent with previous findings using similar tasks with preschoolers.

The Battery we study could introduce new clinical practice that add to clinical observations or questionnaires rigorous measurement tools with good psychometric properties.

Overall, in the following part of the discussion, we consider each of our findings in turn.

The utility of a battery of assessment of FE in children is found both in developmental disorders, and in children with acquired brain injury.

The role of executive functions in developmental psychopathology has been the focus of considerable research and a feature of conceptual models for a range of conditions including, but not limited to ADHD, autism, schizophrenia, bipolar disorder, learning disorders and aggression /conduct problems. (Halperin, 2016).

Furthermore, There is growing recognition that executive functions frequently impaired by traumatic brain injury in children and mediates the neurobehavioral sequelae exhibited by these patients.

The effects of frontal lesions in childhood are not silent and do not differ radically from those seen in adults. Among the syndromes socioemotive observed in adults the most commonly identified is that psychopathic. Deficits in executive functions, related to the behaviour and attention (including undue attention to detail) and integration time, are a prominent feature of children with head trauma. In addition to these serious deficits were observed in children with head trauma lack of interactions with peers, lack of empathy and significant differences in behavioural characteristics premorbid

After describing the structure of and procedures for BAFE it ends with a short summary of the project.

For the selection of the evidence in the international literature for their clinical relevance, the BAFE is a unique tool in Italy for assessing executive function in preschool. The battery is based on a concept, credited by the main research: a representation articulated in a profile of neuropsychological processes relatively independent.

The battery was built to describe the mechanisms of the various sub-domains of FE and to examine their maturation and their modification by specific and targeted rehabilitation training.

Beyond the borders of the academic research community, as evidence grows on the importance of EF in early childhood, clinicians and educators need valid diagnostic and assessment tools (e.g. to predict school readiness, etc). The present study introduces a promising new EF battery that is sensitive to age differences among preschoolers from 3 to 6 years.

The construction of this diagnostic tool fits into this perspective and is a tentative to start a reflection shared between all operators who deal with neurodevelopment.

REFERENCES

- Baddeley, A. (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath. *Psychonomic Bulletin & Review*, 7(3), 544-549.
- Baddeley, A., & Wilson, B. A. (2002). Prose recall and amnesia: implications for the structure of working memory. *Neuropsychologia*, 40(10), 1737-1743.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65.
- Benes, F. M. (2001). The development of prefrontal cortex: The maturation of neurotransmitter systems and their interactions. *Handbook of Developmental Cognitive Neuroscience*, 2, 79-92.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238.
- Bentler, P. M., & Bonnet, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136-136
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental neuropsychology*, 28(2), 595-616.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032-1053.
- Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, 87(4), 299-319.
- Cohen, J. D., Aston-Jones, G., & Gilzenrat, M. S. (2004). A Systems-Level Perspective on Attention and Cognitive Control: Guided Activation, Adaptive Gating, Conflict Monitoring, and Exploitation versus Exploration. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 71-90). New York: Guilford Press.
- Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLoS Biology*, 10(3), e1001293.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135-168.
- Diamond, A. (2014). Want to optimize executive functions and academic outcomes?: simple, just nourish the human spirit. In In Sera M. D.,

- Zelazo P. D. (Eds.), *Minnesota Symposia on Child Psychology* (Vol. 37, p. 205). NIH Public Access.
- Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to “Do as I say, not as I do”. *Developmental Psychobiology*, 29(4), 315-334.
- Dunn, L., & Dunn, L. (1981). *PPVT-R Manual*. Circle Pines, MN: American Guidance Service. (2000) Italian version Stella G., Pizzoli C., Tressoldi P.E.. Torino: Omega Edizioni
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466.
- Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, 52(2), 326.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10(4), 483-527.
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin*, 134(1), 31.
- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children 312–7 years old on a stroop-like day-night test. *Cognition*, 53(2), 129-153.
- Halperin, J. M. (2016). Executive functioning—a key construct for understanding developmental psychopathology or a ‘catch-all’ term in need of some rethinking?. *Journal of Child Psychology and Psychiatry*, 57(4), 443-445.
- Hodapp, A. F. (1993). Correlation between Stanford-Binet IV and PPVT—R Scores for Young Children. *Psychological Reports*, 73(3\_suppl), 1152-1154
- Howard, K., Anderson, P. J., & Taylor, H. G. (2010). Executive functioning and attention in children born preterm. In Vicki Anderson, Rani Jacobs, Peter J. Anderson (Eds.), *Executive Functions and the Frontal Lobes: A Lifespan Perspective*. United Kingdom, Hove: Psychology Press.
- Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, 16(2), 233-253.
- Hughes, C., & Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and

- internalizing behaviors and self-perceived academic success at 6 years of age. *Journal of Experimental Child Psychology*, 108(3), 663-676.
- Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive function in preschool children: Examination through everyday behavior. *Developmental Neuropsychology*, 26(1), 403-422.
- Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality*, 71(6), 1087-1112.
- Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child development*, 67(2), 490-507.
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY – Second edition (NEPSY II)*. Chicago: Psychological Corporation
- Lezak, M. D. (2004). *Neuropsychological assessment*. USA: Oxford University Press.
- Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four-to eight-year-old children. *Neuropsychologia*, 36(3), 273-293.
- Lunt, L., Bramham, J., Morris, R. G., Bullock, P. R., Selway, R. P., Xenitidis, K., & David, A. S. (2012). Prefrontal cortex dysfunction and 'Jumping to Conclusions': Bias or deficit?. *Journal of Neuropsychology*, 6(1), 65-78.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167-202.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., & Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, 108(7), 2693-2698.
- Mueller, U., Kerns, K. A., & Konkin, K. (2012). Test-retest reliability and practice effects of executive function tasks in preschool children. *The Clinical Neuropsychologist*, 26(2), 271-287.
- Muthen, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class-latent growth modeling. In L. M. Collins &

- A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 291-322). Washington, DC: American Psychological Association.
- Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). Unpublished Technical Report.
- Muthén, L. K., & Muthén, B. O. (2004). *Mplus User's Guide: Statistical Analysis with Latent Variables: User's Guide*. Muthén & Muthén.
- Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 241-268.
- Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 1*(4), 349-370.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of child psychology and psychiatry, 37*(1), 51-87.
- Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly (1982)*, 131-147.
- Roid, G. H., & Miller, L. J. (1997). *Leiter international performance scale-revised: Examiners manual*. Wood Dale, IL: Stoelting.
- Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102*(41), 14931-14936.
- Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 298*(1089), 199-209.
- Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N. E. E. A., & Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature, 440*(7084), 676-679.
- Sonuga-Barke, E. J. (2003). The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *Neuroscience & Biobehavioral Reviews, 27*(7), 593-604.
- Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition, 20*(1), 8-23.
- Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(10), 1231-1238.

- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10.
- Urgesi, C., Campanella, F., & Fabbro, F. (2011). NEPSY-II. *Contributo alla Taratura Italiana, Giunti OS, Firenze*.
- Valeri G., Stievano P., Ferretti M.L., Mariani E., Pieretti E. (2015) *BAFE Batteria per l'Assessment delle Funzioni Esecutive*, Firenze: Hogrefe Editore.
- Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, 108(3), 436-452.
- Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences*, 8(1), 12-17.
- Zelazo, P. D., & Müller, U. (2002). The balance beam in the balance: Reflections on rules, relational complexity, and developmental processes. *Journal of Experimental Child Psychology*, 81(4), 458-465.
- Zelazo, P. D., Qu, L., Müller, U., Schneider, W., Schumann-Hengsteler, R., & Sodian, B. (2005). Hot and cool aspects of executive function: Relations in early development. In Schneider, W., Schumann-Hengsteler, R., & Sodian, B. (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (71-93). United Kingdom, Hove: Psychology Press.