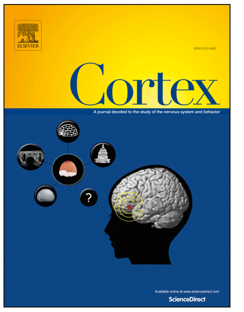
Journal Pre-proof



Neuro-cognitive architecture of executive functions: A latent variable analysis

Ettore Ambrosini, Sandra Arbula, Chiara Rossato, Valentina Pacella, Antonino Vallesi

PII: S0010-9452(19)30281-3

DOI: https://doi.org/10.1016/j.cortex.2019.07.013

Reference: CORTEX 2692

To appear in: *Cortex*

Received Date: 6 March 2019

Revised Date: 15 May 2019

Accepted Date: 25 July 2019

Please cite this article as: Ambrosini E, Arbula S, Rossato C, Pacella V, Vallesi A, Neuro-cognitive architecture of executive functions: A latent variable analysis, *CORTEX*, https://doi.org/10.1016/j.cortex.2019.07.013.

Neuro-cognitive architecture of executive functions: A latent variable analysis

Ettore Ambrosini[1,2,#], Sandra Arbula[3], Chiara Rossato[3], Valentina Pacella[2,4], Antonino Vallesi[1,5,#]

[1] Department of Neuroscience & Padova Neuroscience Center, University of Padova, Italy

[2] Department of General Psychology, University of Padova, Italy

[3] Department of Neuroscience, University of Padova, Italy

[4] Social and Cognitive Neuroscience Laboratory, Department of Psychology, Sapienza University of Rome, Italy.

[5] Brain Imaging and Neural Dynamics Research Group, IRCCS San Camillo Hospital, Venice, Italy

[#] Correspondence should be addressed to:

- Ettore Ambrosini, Department of Neuroscience, University of Padua, Via Giustiniani, 5, 35128 Padova, Italy. E-mail: ettore.ambrosini@unipd.it
- Antonino Vallesi, Department of Neuroscience, University of Padua, Via Giustiniani, 5, 35128 Padova, Italy. E-mail: antonino.vallesi@unipd.it

Competing interest statement

The authors have no competing interests to declare.

Abstract

Executive functions refer to high-level cognitive processes that, by operating on lower-level mental processes, flexibly regulate and control our thoughts and goal-directed behavior. Despite their crucial role, the study of the nature and organization of executive functions still faces inherent difficulties. Moreover, most executive function models put under test until now are brain-free models: they are defined and discussed without assumptions regarding the neural bases of executive functions. By using a latent variable approach, here we tested a brain-centered model of executive function organization proposing that two distinct domain-general executive functions, namely, criterion setting and monitoring, may be dissociable both functionally and anatomically, with a left vs. right hemispheric preference of prefrontal cortex and related neural networks, respectively. To this end, we tested a sample of healthy participants on a battery of computerized tasks assessing criterion setting and monitoring processes and involving diverse task domains, including the verbal and visuospatial ones, which are well-known to be lateralized. By doing this, we were able to specifically assess the influence of these task domains on the organization of executive functions and to directly contrast a process-based model of EF organization versus both a purely domain-based model and a process-based, but domain-dependent one. The results of confirmatory factor analyses showed that a purely process-based model reliably provided a better fit to the observed data as compared to alternative models, supporting the specific theoretical model that fractionates a subset of executive functions into criterion setting and monitoring with hemispheric specializations emerging regardless of the task domain.

Keywords: executive functions, monitoring, task-setting, latent variable analysis, factor analysis, hemispheric specialization.

Introduction

Executive functions (EFs) refer to high-level cognitive processes that, by operating on lower-level mental processes, flexibly regulate and control our thoughts and behaviors enabling us to achieve internally represented goals (Koechlin, Ody, & Kouneiher, 2003; Miller & Cohen, 2001; Stuss, Shallice, Alexander, & Picton, 1995). EFs, also referred to as cognitive – or executive – control, are critically important in novel or cognitively demanding circumstances, that is, when one has to "concentrate and pay attention" (Diamond, 2013). They are also needed in various situations where task goals have to be actively maintained and the processing of task-relevant information has to be boosted in spite of internal and external distracting information, or when habitual or prepotent response tendencies must be overcome (MacDonald, Cohen, Stenger, & Carter, 2000; Burgess & Shallice, 1996; Baddeley, 1996).

Despite the crucial role of EFs in human behavior, and despite the extensive efforts made by various scholars in the last years, the study of the nature and organization of EFs still faces inherent difficulties. The first of these difficulties consists in the lack of a univocal definition and an explicit and clear operational characterization of EFs. Tens of definitions have been proposed for EFs over the past years, and just as many distinct mental processes have been attributed to EFs (Barkley, 2012; McCloskey & Perkins, 2013; Banich, 2009). EFs would thus be best conceptualized not as a unitary construct, but rather as a meta-construct or, better, as a multifaceted phenomenon (Eslinger, 1996; Jurado & Rosselli, 2007; MacPherson et al., 2019; Miyake & Friedman, 2012).

Another difficulty faced in the study of EFs is their accurate, reliable, and valid assessment: EFs are not only difficult to define, but also challenging to assess and quantify (Chan et al., 2008). The reasons for this difficulty are numerous, including the low internal or test-retest reliability of available measurement instruments (Miyake et al., 2000; Phillips, 1999) and, especially, the so-called task-impurity problem (Burgess, 1997; Phillips, 1999; Weiskrantz, 1992). In other words, EF tasks are almost never process-pure, that is, they do not measure exclusively the process or function they were supposed to assess. On the contrary, any score derived from an EF task unavoidably contains an (often substantial) amount of variance that is not directly related with the EF of interest, thus making it problematic to isolate specific EF components. Task impurity of EF paradigms is inevitable because, by definition, EFs manifest themselves by acting on other non-executive cognitive processes that are not central to the target EF but nonetheless could influence performance in a given task. What is worse, the specific influence of non-EF processes to the performance in EF tasks fluctuates greatly across tasks as well as across individuals, a point on which we will return later.

One approach that is especially helpful for limiting the impact of these issues is to employ hypothesis testing-based methods such as latent variable analysis, or confirmatory factor analysis (CFA), which consists in testing theoretical models of the interplay between different EF processes or, in other words, in assessing whether and to what degree they interact with one another. Latent variables, indeed, would exclusively capture the portion of variance that is in common across multiple EF measures and, by definition, this common variance does not include task-specific variance or random measurement errors (e.g., Miyake et al., 2000).

The use of this approach is now common to investigate the interplay between different EFs and, thus, to test specific theoretical models of EF organization (e.g., Miyake et al., 2000; Fournier-Vicente et al., 2008; Rau et al., 2016); moreover, this approach has been proven to be very helpful to unveil how EFs are related to other higher-order capacities such as working memory and fluid intelligence (e.g., Miyake et al., 2001; Friedman et al., 2006; Unsworth et al., 2009; Dang et al., 2014; Wongupparaj et al., 2015) or, as another example, to investigate the EF role in mediating age-related cognitive decline (e.g., Salthouse et al., 2003). However, most of the recent works employing the latent variable approach to investigate the organization of EFs on healthy individuals share a potential limitation. Specifically, although the EF construct has been mainly used until very recently to indicate high-level cognitive processes mainly implemented in prefrontal cortex (PFC) and, more generally, in fronto-parietal networks, most of the EF models put under test in previous latent variable studies are “brain-free” models. In other words, they were created without assumptions regarding how the brain mediates EFs, or their neural basis; thus, they somewhat neglect the strong link between the PFC function and the organization of EFs that has characterized the study of EFs since its origin (but see, for example Bettcher et al., 2016).

Moreover, the investigation of EF organization has mostly neglected an aspect that plays an important role in determining the inter-individual variability in EF performance, and that can contribute to reveal the organization of EFs (see for example Postle et al., 2000; Park et al., 2006), that is, the influence of the different domains of functioning. Indeed, most of the existing studies did not manipulate or control in a systematic way the task domains involved (or the kind of material used) in the different EF tasks employed to explore the EFs latent structure, even though it is well acknowledged that there is considerable inter-individual variability as individuals tend to have areas of relative strength and weakness in specific task domains (e.g., verbal, spatial, numerical) and lower-level sensory-motor processing, which may affect their performance in experimental tasks tapping EFs (Deary, Penke & Johnson, 2010; Fillmore, Kempler & Wang, 2014; Naber et al., 2016).

Here we tackle these issues by examining a theoretical model of EFs that is strongly based on the organization of PFC (Stuss, 2011; Shallice, Stuss, Picton, Alexander, & Gillingham, 2007, 2008; Stuss & Alexander, 2007; Vallesi, 2012), as it was based on a large number of neuropsychological, electrophysiological and neuroimaging studies. In particular, this EF model proposes that two distinct domain-general EFs may be dissociable not only functionally, but also anatomically, with a left-right hemispheric specialization of PFC and related neural networks for, respectively, the criterion (or task) setting and the monitoring processes. In particular, criterion setting can be defined as the cognitive control function in charge of forming and/or selecting associations or rules that are relevant for accomplishing a given task and achieving a goal, as well as actively suppressing the interfering, task-irrelevant ones (Stuss & Alexander, 2007; Vallesi, McIntosh, Crescentini, & Stuss, 2012; Fletcher, Shallice & Dolan, 2000; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Vallesi, McIntosh, Alexander & Stuss, 2009; Vallesi, 2012; see Mostofsky & Simmods, 2008, for the alternative proposal that this function is mainly implemented by dorso-medial prefrontal regions). Monitoring can be instead defined as the cognitive control function in charge of actively maintaining representations of task-relevant goals and events and checking and examining their relative status in relation to each other and to the flow of events, in order to make behavioral adjustments and to optimize performance when needed (Petrides, 2005; Stuss & Alexander, 2007; Vallesi, 2012).

To this aim, we asked a sample of healthy participants to perform a battery of computerized tasks assessing criterion setting and monitoring. Importantly, different versions of most of these tasks were designed, so that the distinct EFs were assessed by using materials and/or tasks that rely on low-level cognitive processes that are well known to be lateralized, such as verbal (left-lateralized) and visuospatial and implicit temporal (right-lateralized) processes. By doing this, we were able to specifically assess the influence of these task domains on the organization of EFs and to directly contrast a process-based, domain-independent model of organization of EFs – in which the executive scores are explained by two latent variables representing the criterion setting and monitoring constructs – with a purely domain-based model – in which the executive scores are explained by two latent variables representing left- and right-lateralized low-level cognitive processes involved in the tasks we used – but also with a process-based, but domain-dependent model, in which the executive scores are explained by four latent variables representing the interaction between the two lateralized EF constructs and the two lateralized task domains.

Materials and Methods

No part of the study analyses was pre-registered prior to the research being conducted. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Participants

One-hundred fifty-seven young participants (98 F; mean age = 22.9 years, range = [19 33] years; mean education level = 16.7 years, range = [13-21]) took part in the study voluntarily after providing their written, informed consent. Most of them were enrolled in university courses tapping different disciplines and were recruited from the relative classes or through web-based advertisement and word of mouth from the authors' networks.

The sample size was chosen so to have more than ten observations for each parameter (Bentler, 1995) and so to be conservatively greater than that recommended for a variables-to-factor ratio of 6, a two-factor solution, a low communality level, and a good-level criterion (Mundfrom, Shaw, & Ke, 2005). All participants reported having normal or corrected-to-normal visual acuity and normal color vision. None of the participants reported having any history of, or current psychiatric or neurological conditions. The study was approved by the Bioethical Committee of the Azienda Ospedaliera of Padova and was conducted according to the guidelines of the Declaration of Helsinki (World Medical Association, 2013). According to the scores of the

Edinburgh Handedness Inventory (Oldfield, 1971), which was administered at the beginning of the first experimental session, fourteen participants were left-handed[1].

General Procedure

The study was divided in two separate experimental sessions. During the first one, a group of five tasks was first administered in a balanced order among the participants (i.e., the foreperiod task, the two Stroop tasks, the color-shape task-switching, and the dichotic listening task)[2]. In the second session, participants performed a paradigm that we created ad-hoc to investigate both the task-switching/criterion setting abilities and the monitoring ones in the verbal and spatial domains by using the same materials and trial structure, which we called 'criterion-setting + monitoring' (CSM) tasks. These CSM tasks were used and validated in a number of studies from our research group in the last years (e.g., Ambrosini & Vallesi, 2016, 2017; Capizzi, Ambrosini, Arbula, Mazzonetto, & Vallesi, 2016a, 2016b; Vallesi, Arbula, Capizzi, Causin, & D'Avella, 2015).

Participants were tested in a quiet and normally illuminated room. They were seated in front of a 17" computer screen (refresh rate: 60 Hz, resolution: 1366 × 768) at a distance of approximately 60 cm.

Behavioral Tasks, Procedures and Data Preparation

No part of the study procedures was pre-registered prior to the research being conducted.

Verbal and spatial criterion setting + monitoring (CSM) tasks

The CSM paradigm was a revised version of the task used in our previous studies (Capizzi et al., 2016a, 2016b; Vallesi et al., 2015), and consisted of verbal and spatial task-switching and

[1] Note that control analyses performed after excluding these participants confirmed the results reported here.

[2] In a sub-group of participants (*n* = 76), four memory tasks (namely, the matrix/symmetry span and the letter/operation span; Unsworth, Heitz, Schrock, & Engle, 2005; Foster et al., 2015) were administered after the first five tasks. For the sake of completeness, we note here that the results of the analyses performed in this sub-sample of participants including the memory scores did not modify the conclusions we draw in the present study.

monitoring sessions (Figures 1A and 1B, respectively), that were administered in a counterbalanced order.

Stimuli were 36 words: 18 proper nouns (9 males and 9 females) and 18 common nouns (9 males and 9 females). The proper nouns were personal names (e.g. “luca”) and names of States (e.g. “cina”, the Italian word for China), while the common names consisted of non-living things (e.g. “taxi”) and generic terms referring to people (e.g. “sposa”, the Italian word for bride). All the words were created with 3D effects and 3D rotation allowing them to assume a clockwise or a counterclockwise rotation (i.e., roll) and an upward or downward rotation (i.e., pitch). In the monitoring, non-monitoring, and single-task conditions (see below), each word was colored in either green or brown, but the word color was not relevant for the task. On the contrary, each word was colored in either blue or red throughout the task-switching condition, as the color cued the task to be performed (see below). All words were written in lowercase letters, subtending on average 3.2 degrees of visual angle (Fig. 1A-B; for a more detailed description of the stimulus features, see Capizzi et al., 2016a, 2016b; Vallesi et al., 2015).

-----*Insert Figure 1 about here*-----

The verbal task was subdivided in the gender- and name-type subtasks. In these sub-tasks, respectively, participants were required to press the “f” key of the computer keyboard with the index finger of the left hand if the word was either a female (gender) or a proper (name) noun, while they had to press the “k” key with the index finger of the right hand if the word was either a male (gender) or a common (name) noun. The spatial task was subdivided in the roll and pitch subtasks. In these spatial subtasks, participants were required to press the “f” key when the word was either oriented counterclockwise (roll) or rotated downward (pitch), whereas participants were instructed to press the “k” key if the word was either oriented clockwise (roll) or rotated upward (pitch). For both the spatial and the verbal tasks, the mapping of categories to response keys was counterbalanced across participants.

Each of the verbal and spatial monitoring sub-sessions included two non-monitoring and two monitoring blocks (each comprising 32 trials), both involving gender and name (or pitch and roll) subtasks, which were presented according to the following order: non-monitoring subtask 1, monitoring subtask 1, non-monitoring subtask 2 and monitoring subtask 2.

During the verbal and spatial non-monitoring blocks, participants had to simply perform one of the corresponding subtasks; instead, during the verbal and spatial monitoring blocks, respectively, participants had to perform not only the corresponding subtasks but, at the same time, they also had to monitor for the occurrence of target words, that is, words that contained the letter “v” (verbal monitoring task) or were rotated exactly by 45° in either direction (spatial monitoring task). If this was the case, they had to interrupt the verbal/spatial subtask and to press the spacebar.

Each of the verbal and spatial task-switching sub-sessions included two single-task and four mixed-task blocks each comprising 32 trials, both involving gender and name (or pitch and roll) sub-tasks, which were presented according to the following order: single-task 1, single-task 2, mixed-task 1 to 4.

During the verbal and spatial single-task blocks, participants had to simply perform one of the corresponding subtasks; instead, during the verbal and spatial mixed-task blocks, respectively, the color of the word instructed participants about the specific verbal or spatial sub-task they had to perform on any given trial. The blue and red colors were associated, respectively, with the name and roll subtasks and with the gender and pitch sub-tasks. The subtasks were thus either repeated (repeat trials) or switched (switch trials) from trial to trial unpredictably based on a pseudorandom sequence.

Each experimental trial started with the presentation of a 400 ms blank grey screen, which contained a grey frame lighter than the background color. Then, the word was presented inside the frame, for 2000 ms. The ITI lasted 1400 ms, during which the blank screen was again displayed.

At the beginning of the experimental sessions, participants practiced the spatial and verbal tasks in every condition. Each block was made up of 10 trials and, at the end of each trial, participants received a feedback message according to their performance (the Italian word for "wrong" written in red, or the Italian word for "good" displayed in blue).

The response times (RTs) of the spatial and verbal blocks were computed separately. Error and post-error trials (10.7% of trials), anticipations (RTs < 120 ms, < 0.1% of trials), trials with missed responses (2.3% of trials), and target-word trials in the monitoring blocks were not taken in account for the analysis. An inverse transformation (1000/RTs) was then computed on the RTs to improve normality. Next, in order to obtain central tendency measures of participants' performance that were as robust as possible against the influence of outliers, we computed for each participant and monitoring condition a robust M-estimator of location (robust mean RTs, see Ambrosini & Vallesi, 2016). This estimation procedure uses the logistic psi-function and the median absolute deviation as the auxiliary scale estimate, as implemented by the *mloclogist* and *madc* functions in the LIBRA Matlab library (Verboven & Hubert, 2005, 2010), and it is robust to non-normality and sample size (Rousseeuw & Verboven, 2002). From the difference between the performance in the non-monitoring blocks and that for non-target trials of the monitoring blocks, we estimated the monitoring effects of the spatial and verbal tasks, which are thought to depend on the activity of right-lateralized fronto-parietal regions mediating monitoring processes (Vallesi, 2012; see also Capizzi et al., 2016b).

For the task-switching paradigms, each participant's RTs from single-task, repeat and switch trials were inverse-transformed and their robust mean was computed as described above. Then, we obtained the switching and mixing costs for the spatial and verbal tasks as, respectively, the difference in performance between switch and repeat trials and between repeat and single trials. Switching and mixing costs are thought to depend, respectively, on the left-lateralized criterion setting and on the right lateralized monitoring processes (Ambrosini & Vallesi, 2016; Capizzi et al., 2016a). Moreover, a number of imaging studies has shown differential activation patterns in lateral PFC related to cognitive control processes mediating switching and mixing costs (Badre & Wagner, 2006; Braver et al., 2003; Wang et al., 2009; Kim et al., 2011a; Dreher et al., 2002).

Color-shape task-switching task

The task was adapted from Babcock and Vallesi (2015). Stimuli were heart and star shapes (visual angle: 2.1°by 2.1° and 1.8 by 1.8°, respectively) colored in red or blue, which were presented individually at the center of the screen (see Figure 1D). Participants had to respond according to either the shape or the color of the presented stimulus, respectively, for the shape and color tasks. The type of task to be performed was indicated by a cue (3.8° by .9° of visual angle) which was visually presented above the stimulus, and was composed of either three small black figures (circle, rectangle and square) for the shape task, or three small colored rectangles (purple, orange and yellow) for the color task. A fixation cross was presented at the beginning of each trial for 1500 ms and was followed by the presentation of the cue. The elapsing time between the cue presentation and the target occurrence (i.e., the cue-to-target-Interval) lasted either 100 or 1000 ms (only the 100 ms long CTI was included in the analysis, because it is timing-wise more comparable to the switching costs observed in the CSM task, where the cue was embedded within the target). The task stimulus was presented below the cue, which remained on the screen until the participant's response (see Figure 1D). The trial ended when participants gave their response. The incorrect responses were followed by a short auditory error feedback.

The paradigm was arranged in three blocks. In two single-task blocks, participants were required to perform two types of subtasks, one at a time in different blocks. In the "shape" task, participants were required to distinguish the star from the heart by pressing either the left or the right pointing arrow of the keyboard. On the opposite, in the "color" single task they had to categorize the stimuli according to their color. These blocks were each composed by 6 practice trials and 24 test trials (single-task trials).

During the third block, which was named mixed-task block, the categorization rule changed unpredictably trial by trial, as indicated by the cue. The task was either repeated (repeat trials) or switched (switch trials) from trial to trial based on a pseudorandom sequence. The four possible response-to-button mappings (1. left: red/heart, right: blue/star; 2. left: red/star, right: blue/heart; 3. left: blue/heart, right: red/star; 4. left: blue/star, right: red/heart) were counterbalanced across participants. This block was composed of 10 practice trials and 192 test trials divided into four sub-blocks of 48 trials.

The RTs for single-task, switch and repeat trials were taken into account in the analysis, and error and post-error trials (9.2% of trials), as well as first trials of every (sub-)blocks and trials with RTs faster than 120 ms (<0.1% of trials) were discarded. After the inverse transformation, the robust mean RTs were computed for each subject and condition. Two indices were finally estimated: the switch cost (the performance difference between switch and repeat trials) and the mixing cost (the performance difference between repeat and single-task trials), which respectively measure criterion setting and monitoring abilities (Ambrosini & Vallesi, 2016).

Verbal and spatial Stroop tasks

Both verbal and spatial versions of the Stroop task were adapted from previous studies from our laboratory (Puccioni, Vallesi, 2012a; Puccioni & Vallesi, 2012b; see Figure 1E). For both tasks, a practice block of 16 trials preceded the test phase that was subdivided into two blocks of 64 trials each.

The stimuli of the verbal Stroop task were the Italian words for yellow (GIALLO), red (ROSSO), blue (BLU) and green (VERDE). The words subtended on average 2.8° x 0.7° of visual angle and were presented in four ink colors: yellow, red, blue and green. In the congruent condition, the ink color corresponded to the meaning of the words, whereas in the incongruent condition the ink color and the words meaning did not match. The two conditions were intermixed trial by trial in a pseudorandom order so to have no repetitions of either color or meaning of the word on subsequent trials, thus minimizing both positive and negative priming confounds (see Puccioni & Vallesi, 2012b, for details). The ink colors were mapped on four keys of the keyboard in two counterbalanced orders (from left to right: 1. blue, red, green, and yellow; or 2. yellow, green, red, and blue). Participants were required to press the key corresponding to word ink colors, ignoring word meanings. The stimuli appeared at the center of the screen for 500 ms and were followed by a blank screen of 2000 ms. The response acquisition time lasted 2500 ms, starting with the stimulus onset, until the end of the blank. The inter-trial interval (ITI) varied randomly between 250 and 700 ms (see Figure 1E).

During the spatial Stroop task, participants were required to fixate a cross that appeared at the center of the screen. The target stimuli were four arrows (2.5° x 1.5° of visual angle) that pointed toward north-east, north-west, south-east and south-west. The arrows were presented one at time and could be positioned in correspondence with one of the corners of the screen (upper right, upper left, lower right, or lower left), at a distance of 8 cm from the fixation cross. As in the verbal version of the task, two conditions were implemented: in the congruent condition, the arrow position on the screen corresponded to its orientation (e.g., it appeared in the upper right corner pointing to north-east); in the incongruent position, the arrow had a different orientation than its position on the screen (e.g. it appeared in the upper right corner pointing to south-west). As in the verbal version, the two conditions alternated pseudo-randomly in order to avoid position or orientation repetitions and thus minimize the priming effects. The arrow orientations were mapped on four keys of the keyboard (r, v, o, m) that were spatially arranged to reflect the spatial characteristics of the arrow (i.e., its position and orientation); participants were instructed to press the key that corresponded to the arrow orientation (i.e., for example, the r key for the arrow pointing to north-west), while ignoring the position on the screen where the arrow appeared. As in the verbal task, the response acquisition time corresponded to the stimulus appearance time (500 ms) plus the duration of a blank (2000 ms) that followed the stimulus offset. The ITI duration varied randomly between 250 and 700 ms.

The RTs were taken into account for the analysis for both Stroop tasks, separately. Error and post-error trials (14.9% and 9.2% of trials for the verbal and spatial Stroop, respectively) and trials with RTs faster than 120 ms (<0.1% of trials in both cases) were excluded from the analysis. After the inverse transformation, we computed the robust mean RTs for the Congruent and Incongruent conditions, for each subject, as described above. The Stroop effect was calculated as the difference between the performance in the congruent and incongruent conditions. It was hypothesized to depend on the left-lateralized criterion setting process (Ambrosini & Vallesi, 2017). Moreover, a number of meta-analyses of fMRI studies on Stroop performance have shown that differential activation patterns induced by interference resistance processes are commonly found in left prefrontal and inferior parietal regions (Neumann, von Cramon, & Lohmann, 2008;

Laird, et al., 2005; Laird, et al., 2005; Neumann, Lohmann, Derrfuss, & von Cramon, 2005, cf. Cieslik et al., 2015).

Dichotic listening task

A couple of different consonant-vowel syllables were auditorily presented at the same time to the left and the right ear, through the headphones. The syllables were "ka", "ta", "da", "pa", "ga" and "ba". The task included three different attentional instructions that were presented in different blocks: Non-Forced (NF), Forced-Right (FR) and Forced-Left (FL). In the Non-Forced condition, participants were required to report vocally the syllable that was better perceived in each trial, independently of the ear of presentation. The NF condition was always presented in the first block. In the FL and FR conditions, participants were required to orient their attention toward one specific ear (according to the block instructions) and report the syllable that was presented to that ear. The FL and FR conditions were presented in the second and third blocks, and their order of presentation was balanced across subjects. Each block was divided in 30 trials, with an ITI of 3000 ms (Figure 1F).

Trials with missed response (0.2% of trials) were excluded from the analysis. The Attentional Shift Index (ASI; Asbjørnsen & Bryden, 1998) was computed as an unbiased measure of the participants' ability to follow the attention orienting instructions. The ASI was calculated as the sum of the logarithmic transformation of the odds ratio of correct responses (i.e., the responses correctly reported from the attended ear) and intrusions (i.e., the responses reported from the unattended ear) for each of the two forced conditions. Finally, we estimated the standardized attentional shift index (zASI) by using the formula reported by Asbjørnsen & Bryden (1998): $(\ln(RE_{FR}/LE_{FR}) + \ln(LE_{FL}/RE_{FL})) /\mathrm{sqrt}(1/RE_{FR} + 1/RE_{FL} + 1/LE_{FR} + 1/LE_{FL})$, where RE_{FR} and LE_{FR} are the number of correct reports from the right and left ear, respectively, when attention was directed to the right ear (FR); and RE_{FL} and LE_{FL} are the number of correct reports from the right and left ear, respectively, when attention was directed to the left ear (FL). This index assesses the participants' ability to selectively focus on the information coming from the indicated ear and, at the same time, suppress the interfering information coming from the other ear; as such, this measure would depend on the left-lateralized criterion setting processes (Hugdahl et al., 2009).

Foreperiod task

The Foreperiod (FP) paradigm was adapted from Vallesi, Arbula and Bernardis (2014; see Figure 1C). A first block of pure FP trials (4 practice trials and 30 test trials) was used to calculate the FP performance. During this first block, an initial cue ('XX', in yellow, 1.5° x 1.2° of visual angle) was presented on a black display together with an auditory warning stimulus (1500 Hz pure tone). The cue was then replaced by the target (a down-pointing white arrow, with the maximum length and width of 2° of visual angle) at a variable foreperiod of either 3000 or 5000 ms. Participants were required to press the spacebar as soon as the target appeared on the screen. Next, a second block of FP trials under dual-task condition (64 trials) was administered, but the data from this block were discarded from the present analyses, as they are not relevant for our hypotheses. The RTs to the target were taken into account for the analysis. Error (i.e., missed response) trials and trials with anticipations (RTs < 120 ms) or responses during the FP (respectively, 0.01%, 0.01%, and 0.04%) were excluded from the analysis. The RTs were inverse-transformed and the robust mean for each participant and FP condition was computed.

The FP score was then derived from the participants' RTs in the long FP conditions as a measure of their ability to monitor the conditional probability of the stimulus occurrence to optimize the response preparation. We also considered the standard FP effect as a measure of participants' monitoring ability. The results of the analyses performed when using this measure did not alter the conclusions we reported here, but the loadings related to the FP effect were not significant. A closer examination of the characteristics of the FP score (and task) as compared to the other measures (and tasks) we used here to assess monitoring abilities helped us interpreting this result suggesting that, differently from the mixing and monitoring costs, the FP effect would be a relative (and maybe less sensitive) measure of monitoring abilities. Indeed, both the mixing and monitoring costs are computed by subtracting the participants' RTs in a purely baseline condition (i.e., the single-task blocks in the task-switching paradigms and the non-monitoring blocks in the monitoring tasks), which thus does not require the particular involvement of monitoring processes, from the RTs obtained in an experimental condition that is assumed to specifically involve monitoring processes (i.e., respectively, the repeat trials in the mixed-task blocks of the task-switching paradigms and the non-target monitoring trials in the monitoring block of the monitoring task) (see above). On the contrary, the FP effect is computed by subtracting the participants' RTs in the long FP condition from those in the short one, which both involve monitoring processes as defined here, although to different extents. Moreover, the RT increase in the short FP condition is also influenced by non-strategic processes responsible for sequential effects, while the long FP condition is only minimally influenced by these processes (e.g., Drazin, 1961; Steinborn & Langner, 2012; Vallesi, Lozano & Correa, 2013). Thus, the inclusion of the short FP condition in the calculation of monitoring-related effects would make this measure somewhat spurious. These are the reasons why we finally decided to use an alternative measure of the FP performance that would represent a measure of monitoring ability more similar to those derived from the other monitoring tasks included in this study.

The performance in the FP paradigm is thought to rely on the monitoring process (Stuss & Alexander, 2011; Vallesi, 2012). Moreover, converging neuropsychological (Stuss et al., 2005; Vallesi et al., 2007a; also see Triviño, Correa, Arnedo, & Lupiañez, 2010), TMS (Vallesi, Shallice, &

Walsh, 2007b) and fMRI (Vallesi et al., 2009) evidence highlights the involvement of right PFC in monitoring processes mediating the FP performance.

Data Analysis

Data screening prior to the factorial analysis

Due to technical problems, 48 performance scores were not available (corresponding to 2.55% of the total number of scores). For each task, we also excluded from the analysis the data from participants that failed to accurately perform a given task, that is, those who had mean overall RTs more than 2.5 *SD*s slower than the sample mean or accuracy below chance level (n = 21, corresponding to 1.1% of the total scores).

Moreover, since the latent variable analysis is very sensitive to the effect of extreme values and outliers, two trimming procedures were applied to the performance scores. In a first standard outlier analysis, the performance scores were transformed in standardized Z scores and the between-subject score distributions for each behavioral measure was checked, discarding scores more than 2.5 *SD*s away from the sample mean (n = 17, corresponding to .9% of the total scores). These outlier observations and the other missing data were replaced by multivariate imputation by chained equations using the *mice* package (Buuren & Groothuis-Oudshoorn, 2010) in R (R Core Team, 2017). The imputation procedure prevents the reduction in power and the risk of biased estimates (e.g., Graham, 2009; Azur et al., 2011).

After this first step of trimming, the distributions of all of the twelve measures showed acceptable levels of skewness and kurtosis (range = -.07 to .37 and -.82 to .09, respectively). In a second step of the trimming procedure, a bivariate distances analysis was performed to identify bivariate outliers by computing Cook's D values (with a cutoff of 1; Cook & Weisberg, 1982) and Mahalanobis distances (with a cutoff determined using p = .001). None of these statistics reported participants with exceeding values, so no additional observation was removed.

We finally scaled the data and checked them for multicollinearity. All the zero-order correlations were smaller than .41 (in absolute terms, see Zero-order correlations section), all the multiple correlation indices were smaller than .61, and all the squared multiple correlations were smaller than .37, thus indicating little evidence of multicollinearity in the data (Tabachnick & Fidell, 2001).

Confirmatory Factor Analysis

The CFA analyses were performed using the R package "lavaan" (Rosseel, 2012) for latent variable modeling using the maximum likelihood estimation procedure. The dataset and the R script we used are available from our project repository on the Open Science Framework (https://osf.io/tu3nm; see the Readme sheet in the DATAscreening.xlsx file). For each tested model, we estimated the factor loadings and the (co)variances. The indices of fit of the model to the data that we took into account were: the χ^2 statistic, the Root Mean Square Error of Approximation (RMSEA), the Standardized Root Mean Square Residual (SRMR), and the Comparative Fit Index (CFI), which is not overly sensitive to sample size (Friedman & Miyake, 2004). The χ^2 statistic estimates the "badness of fit" of the model, since it measures the difference level between the covariance predicted by the model and the covariance of the observed values, with a small χ^2 value indicating a satisfactory fit. Instead, the other fit indices estimate the "goodness of fit", but their reliability directly depends on a good performance of the χ^2 statistic. The RMSEA is an informative index of the goodness of fit of the model (hypothesizing optimally chosen parameter estimates) with the population covariance matrix. We chose the cutoff value recommended by Hu and Bentler (1999), with values lower than .06 indicating a good fit. For the SRMR, which is the square root of the difference between the observed values covariance matrix and the predicted model covariance matrix, the cutoff value was chosen again based on Hu and

Bentler's (1999) recommendations, with values lower than .08 indicating a good fit. Finally, the CFI compares the goodness of fit between the predicted model and a baseline model; the higher the value of the index is, the better the fit is. The index ranges from 0 to 1 and values higher than .90 commonly indicate an acceptable model fit (e.g., Friedman & Miyake, 2004).

To assess whether a given model was significantly better than another, we performed χ^2 difference tests on nested models. In this test, the χ^2 for the full model is subtracted from that for a nested, restricted model with more degrees of freedom; the difference between the degrees of freedom of the two models is calculated with a similar subtraction. If the resulting χ^2 difference is significant, then the fuller model has a significantly better fit than the nested one. We compared non-nested models by using the Akaike information criterion (AIC), which also takes into account model complexity; the smaller the AIC value is, the better the model fit is. To assess the statistical significance of the factor loadings, we performed one-tailed Wald's *Z* tests. All analyses used an alpha level of .05. All the criteria used for data analysis were established prior to data analysis.

Results and Discussion

Zero-order correlations

As can be seen in Table 1, the zero-order correlations among the twelve measures were generally quite low, in line with the findings from previous individual difference studies of EFs (e.g., Fournier-Vicente et al., 2008; Miyake et al., 2000). However, it is important to highlight that the correlations among the twelve measures were not uniformly low; on the contrary, the correlations among measures chosen to assess the same EF tended to be stronger – and significant – as compared to those among measures considered to assess different EFs, thus showing some degree of convergent and divergent validity.

----Insert Table 1 about here----

Confirmatory factor analysis

Contrasting the domain- versus process-based organization of EFs

The present study aimed to test a theoretical model of EFs that proposes the existence of two domain-general EFs, criterion setting (CS) and monitoring (M), that are dissociable both functionally and anatomically, with a left and right prefrontal gradient of specialization, respectively (Stuss, 2011; Stuss & Alexander, 2007; Vallesi 2012). We aimed to verify the domain-general nature of CS and M by assessing whether, and to what degree, the hypothesized specialization of CS and M was influenced by the left- and right-lateralized low-level processes implicated in EF tasks. We addressed these issues by using CFA analyses, which allowed us to directly contrast the domain-based model (i.e., the Left- vs. Right-lateralized domains – L_R – model) versus the process-based model (i.e., the Criterion Setting vs. Monitoring – CS_M – model) of organization of EFs by evaluating the fit of the corresponding a priori models to the data. The CFA approaches also allowed us to evaluate additional theoretical hypotheses about the organization of EFs by comparing the fit of different alternative models.

<u>*Domain-based models*</u>

We started from a first domain-based measurement model of EFs (L_R model) in which the twelve scores derived from the EF tasks were explained by two factors representing, respectively, the well-known left- and right-lateralization of the low-level cognitive processes involved in the tasks we used: 1) the left-lateralized verbal processes (involved in the verbal version of the CSM and Stroop tasks, as well as in the dichotic listening task) and 2) the right-lateralized temporal and visuospatial temporal processes involved, respectively, in the foreperiod task and in the spatial version of the CSM and Stroop tasks, as well as in the color-shape task-switching task. In constructing this model, it was necessary to allow the error variances deriving from different constituents of each of the task-switching tasks to correlate with each other. Indeed, the covariances between the switching and mixing costs may be due, at least in part, to methodological reasons that cannot be attributable to the latent variables (Friedman & Miyake, 2004).

This model did not fit adequately the data, as evidenced by the fit indices ($\chi^2_{(50)}$ = 104.20, p < .001; AIC = 5276.93; CFI = .643; RMSEA = .083; SRMR = .099) and, more importantly, by the fact that the covariance matrix of latent factors was not positive definite. The inspection of the model fit suggested that the misfit was probably due to the implausibly high covariance between the two latent factors. To verify this impression and to confirm the validity of the loadings obtained by fitting the L_R model, we fitted a similar model that assumes the equivalence of L and R latent factors (L=R model), in which the correlation between Left and Right latent factors was forced to be 1. The estimation of the L=R model parameters was completed without problems, but the fit indices for this model were still poor and very similar to those observed for the L_R model, including a significant chi-square ($\chi^2_{(51)}$ = 106.21, p < .001), which indicates a significant difference between the observed and reproduced covariances, very low CFI (.636), and high RMSEA and SRMR (respectively, .083 and .100). This result, thus, confirmed that the domain-based L_R model fit to the data was inadequate. Moreover, some of the paths between the EF scores and the Left and Right latent variables were not significantly different from zero. Specifically, the path between the zASI and the Left factor (.057, *SE* = .056, Wald's *Z* = 1.027, *p* = .152), and the path between the FP score and the Right factor (.009, *SE* = .064, Wald's *Z* = .144, *p* = .442) were not significant. The latter result could be explained by the fact that the foreperiod task is the only one, among the EF tasks that are considered to involve right-lateralized low-level processes, that implicates the temporal domain instead of the visuospatial domain.

We thus verified whether the exclusion of the FP score would yield a better fit of the resulting domain-based model, which can now be assumed to include a left-lateralized verbal factor and a right-lateralized spatial one (V_S model). Again, this model provided a bad fit, with a covariance matrix of latent factors that was still not positive definite, a significant chi-square ($\chi^2_{(40)} = 93.47$, $p < .001$), very low CFI (.649), and high RMSEA and SRMR (respectively, .092 and .117).[3]

Taken together, these results indicate that the EFs we investigated cannot be adequately fractionated on the basis of the left- and right-lateralized low-level cognitive processes involved in the task we used, which were related either to the specific materials or to the task instructions adopted. Therefore, the great inter-individual variability in the verbal and spatial abilities (Deary et al., 2010; Fillmore et al., 2014) seems to have no significant impact on the organization of EFs. In other words, these results provide support for the hypothesis that EF processes (at least, the ones we investigated here) are not domain-dependent but, rather, are organized in a domain-general way. In order to confirm this hypothesis, we proceeded by testing the alternative, process-based model of the EF organization (see Introduction section).

----*Insert Table 2 about here*----

Process-based models

We thus tested a first process-based measurement model of EFs (CS_M model) that included the twelve scores derived from the EF tasks, which were now explained by two factors representing, respectively, the CS and monitoring M EFs involved in the tasks we used. These EFs

[3] Note that the V=S model in which the correlation between Left and Right latent factors was forced to be 1 provided very similar results to those reported for the unconstrained model.

are considered to rely, respectively, on the activity of the left and right brain hemisphere networks and, in particular, of PFC nodes, regardless of the task domain implicated by the tasks used (Stuss & Alexander, 2007; Vallesi, 2012). Specifically, the CS function is supposed to mediate the performance in the Stroop and dichotic listening tasks, as well as the switching cost derived from the task-switching paradigms. On the contrary, the M function is supposed to mediate the performance in the monitoring tasks included in the CSM tasks, the mixing costs derived from the task-switching tasks, and the FP score derived from the foreperiod task (see Fig. 2).

-----*Insert Figure 2 about here*-----

As can be seen in Table 2, the fit of this model to the observed data was satisfactory. Indeed, the χ^2 statistic was not significant ($\chi^2_{(50)}$ = 60.77, p = .142), the CFI was above the respective .90 cutoff for an acceptable fit (.929), and the RMSEA and SRMR were both below to the respective .06 and .08 cutoff values for an acceptable fit (.037 and .067, respectively). Moreover, the AIC value was lower than that observed for the L_R model (5234 and 5277, respectively), suggesting that this process-based model provided a better fit to the observed data as compared to the domain-based one.

Interestingly, differently from what observed for the L_R model, the path between the zASI and the CS latent factor (.185, *SE* = .092, *Z* = 2.019, *p* = .022) as well as the path between the FP and the M latent factor (.186, *SE* = .101, *Z* = 1.850, *p* = .032) were now significant (Figure 2). This suggests the provisional conclusion that the low loadings we observed above for this measure in the domain-based analysis could have been caused by the low-level domain-dependent characteristics of the tasks used.

Are CS and M unitary or distinct constructs?

One important issue is whether the CS and M latent variables could be truly considered to reflect distinct cognitive control abilities as we hypothesized. As shown in Figure 2, the estimated correlation between CS and M latent factors in the CS_M model was non-significant (.178, SE = .123, Z = 1.450, p = .147), suggesting some degree of independence between CS and M processes. In order to provide further support for this conclusion, we compared the CS_M model to a similar model that assumes the complete independence of CS and M latent factors (unrelated_CS_M model), in which the correlation between CS and M was forced to be 0. As can be seen in Table 2, there is some evidence for preferring the reduced, unrelated_CS_M model over the CS_M one. In fact, even though most of the fit indices were slightly poorer than those of the full CS_M model, the χ^2 difference between the models was not significant ($\Delta\chi^2_{(1)}$ = 1.90, p = .168; Fig. 3). This analysis, thus, suggested that no commonality is shared between CS and M constructs.

-----Insert Figure 3 about here-----

Moreover, the fact that the correlation between CS and M latent variables was far from 1 also supports the idea that the CS and M constructs are indeed separable. This conclusion was further reinforced by an additional analysis that directly compared the CS_M model to an alternative unitary model (CS=M model), which in contrast assumes that CS and M latent variables are in fact the same construct by forcing their correlation to be 1. The unitary CS=M model provided a bad fit to the observed data, with a significant chi-square ($\chi^2_{(51)}$ = 97.11, p < .001), very low CFI, and high RMSEA and SRMR (Table 2). Moreover, the χ^2 difference test showed that the one-factor model provided a significantly worse fit than the full CS_M model ($\Delta\chi^2_{(1)}$ = 36.34, p < .001; Fig. 3), confirming that the CS and M latent variables represent separable constructs.

Are CS and M domain-general?

The results of the CFA analyses examining process-based models confirmed the existence of two distinct cognitive control EFs, the CS and M, that in the theoretical model of EF organization tested here are assumed to be mediated, respectively, by left and right PFC (and relative brain networks) activity irrespective of the task domain and other task specifics (Stuss, 2011; Vallesi, 2012). These findings thus support the idea that EFs are organized as a function of lateralized but domain-general cognitive control processes rather than as a function of a hemispheric specialization of domain-specific processes. The present results are also in line with our previous findings showing a left PFC hemispheric preference for domain-general phasic cognitive control processes mediating Stroop (Ambrosini & Vallesi, 2017) and task-switching (Ambrosini & Vallesi, 2016; Vallesi et al., 2015; Capizzi et al., 2016a) performance and a right PFC hemispheric preference for cognitive control processes mediating monitoring performance (Capizzi et al., 2016b; Tarantino et al., 2017).

Evidence for this conclusion, however, is still not definitive, as other previous findings from our laboratory indicated that hemispheric asymmetries in EFs may be at least partially modulated by the specific characteristics of the tasks to be performed and the relative task domains; in particular, it has been shown that CS processes mediating task-switching (Capizzi et al., 2015; see also Vallesi et al., 2015) and inductive reasoning (Babcock et al., 2015) performance may interact in the brain with the (verbal vs. spatial) domain involved in the task. Therefore, to tackle this issue, we directly examined the fit of a model that assumes the interaction between CS and M cognitive control processes and left- and right-lateralized lower-order cognitive processes (CS_M×L_R model). This model included four latent variables representing, respectively: 1) the CS processes mediating the EF measures derived from tasks involving left-lateralized task domains, that is, the Stroop effect and switching cost derived from the verbal version of the CSM task-switching paradigms, as well as the zASI score from the dichotic listening paradigm; 2) the CS processes mediating the EF measures derived from tasks involving right-lateralized task domains, that is, the Stroop effect and the switching cost derived from the spatial and color-shape task-switching tasks; 3) the M processes measures derived from tasks involving left-lateralized task domains, that is, the monitoring and mixing costs derived from the verbal version of the CSM; and 4) the M processes measures derived from tasks involving right-lateralized task domains, that is, the monitoring and mixing costs derived from the spatial version of the CSM and the color-shape task-switching paradigm, as well as the foreperiod effect. At first sight, the fit of this model to the observed data seemed to be satisfactory (see Table 2). Indeed, the χ^2 statistic was not significant ($\chi^2_{(45)}$ = 54.15, *p* = .165), the RMSEA and SRMR were sufficiently low, and the CFI was above the cutoff for an acceptable fit. However, the χ^2 difference test was far from the significance level ($\Delta\chi^2_{(5)}$ = 6.62, *p* = .250, Figure 3), thus showing that the CS_M×L_R model did not provide a significantly better fit as compared to the CS_M one, which thus has to be preferred over the more complex CS_M×L_R model. More importantly, the covariance matrix of latent factors for the CS_M×L_R model was not positive definite, indicating that the model was not adequate to fit the observed data. The inspection of the model fit suggested that the misfit was probably due to the implausibly high covariance between the left- and right-lateralized task domains for both CS and M latent factors. These results thus provide further support for the conclusion stated above that EFs are organized as a function of lateralized but domain-general CS and M processes.

General Discussion

The present study aimed to test a brain-centered model of EF organization that proposes the existence of two distinct domain-general EFs: criterion setting (CS) and monitoring (M) (Stuss, 2011; Shallice et al., 2008; Stuss & Alexander, 2007; Vallesi, 2012). According to this model, there is a clear anatomical/functional relationship in the organization of the PFC, with a left-right hemispheric preference of PFC and related neural networks for, respectively, the CS and M processes, irrespective of the task domain. To this end, we asked a sample of healthy participants to perform a battery of computerized tasks assessing CS (task-switching – for the switching costs – and interference resistance tasks) and M (task-switching – for the mixing costs – and monitoring tasks) processes. Importantly, these tasks involved task domains that are well-known to be lateralized, such as the verbal (left lateralized) and visuospatial (right lateralized) domains. In this way, we were able to explicitly quantify the influence of these task domains on the organization of EFs by using the latent variable analysis approach.

The results of the CFA analyses showed that a purely process-based model reliably provided a better fit to the observed data as compared to a purely domain-based model that, moreover, was not adequate in fitting the observed data. This finding thus indicates that the two EFs under examination in the present study are organized as a function of lateralized but domain-general cognitive control processes, rather than as a function of a hemispheric specialization of domain-specific processes. Furthermore, in line with this assertion, the CFA analyses provided strong evidence for the diversity of CS and M constructs, corroborating the hypothesis that they are distinct, unrelated EFs. Indeed, the comparison between the CS_M model and a similar model that assumed the complete independence of CS and M latent factors (unrelated_CS_M model) indicated that no commonality is shared between CS and M constructs.

However, one could argue that, even if we assume that higher-order cognitive processes and (in particular) EFs are not organized strictly on the basis of task domains, it is still possible that lower-level processes may at least in part have an impact on their functioning and organization. It is indeed conceivable that the inter-individual variability in EF functioning, which has been shown to be related to differences in task-evoked brain activity (Kim et al., 2011b) and even brain structural organization (Gold et al., 2010; Vallesi et al., 2016), could also be influenced by the inter-individual variability in domain-specific abilities. Indeed, as noted in the Introduction, individuals tend to have areas of relative strength and weakness in terms of verbal, visuospatial, or numerical abilities (Deary et al., 2010; Fillmore et al., 2014), and this could impact not only the way they deal with everyday simple problems and tasks, or the way they learn (e.g., Mayer & Massa, 2003), but also their performance in novel, interfering or complex situations or, in other words, their EF abilities (e.g., Naber et al., 2016). In line with this hypothesis, indeed, some of our previous findings showed that cognitive control processes mediating inductive reasoning (Babcock et al., 2015) and task-switching (Capizzi et al., 2015; see also Vallesi et al., 2015) performance may interact in the brain with the (verbal vs. spatial) domain involved in the task.

Thus, in an additional CFA analysis, we directly verified the hypothesis of an interaction between cognitive control processes and domains. This analysis revealed that the purely process-based model had a better fit as compared to a mixed model that assumed the interaction between CS and M cognitive control processes and left- and right-lateralized lower-order cognitive processes. This result suggests that the specific characteristics of the tasks to be performed and the relative task domains had no influence in determining the functional organization of CS and M processes. The present findings, thus, confirm the process-based theoretical model of EF organization tested here (Stuss, 2011; Vallesi, 2012), indicating the existence of two distinct domain-general cognitive control functions, the CS and M, that would be mediated, respectively, by left and right PFC (and relative brain networks), irrespective of the task domain. Moreover, they confirm and extend our previous findings that showed that CS processes mediating Stroop and task-switching performance share the same neural underpinnings (i.e., left lateral PFC activity) regardless of the task domain (Ambrosini & Vallesi, 2016, 2017; Vallesi et al., 2015; Capizzi et al., 2016a), as well as those that showed that M processes mediating performance in different monitoring tasks share the same neural underpinnings (i.e., right PFC activity) regardless of the task domain involved in the task (Capizzi et al., 2016b; Tarantino et al., 2017; also see Fleck et al., 2006). Taken together, thus, the present results suggest that the high inter-individual variability observed in the verbal and spatial abilities (Deary et al., 2010; Fillmore et al., 2014) would not have a significant impact on how EFs are organized in our brain. This provides support for the hypothesis that the EF processes (at least those investigated here) are not domain-dependent but, rather, are organized in a domain-general way.

Although our tested model was inspired by findings from previous cognitive neuroscience studies, as a future development of this research line, it would be advisable to more directly relate the latent variables underlying behavioral performance on a similar test battery to functional neuroimaging data acquired within the same group of participants during task performance. However, the feasibility of this approach would need to be carefully assessed (e.g., multiple scanning sessions would probably be required, as it takes several hours to perform the whole task battery with an adequate number of trials per condition).

Conclusions

By using a latent variable approach, we provided behavioral evidence supporting the specific theoretical model that fractionates two key executive functions into criterion setting and monitoring, two distinct domain-general executive functions that are supposed to have equally distinct (at least as gradients) neural bases: the activity of the left and right PFC and functionally connected brain networks, respectively. The advantage of the model we examined is represented by its strong foundation on an extensive body of neuropsychological, electroencephalographic and functional neuroimaging studies, in line with the strong emphasis on the brain that has characterized the study of executive functions since its origin.

The presented study represents only a first step. Since it aimed to test a specific theoretical model of the organization of executive functions, our choice of the tasks to analyze and the target latent variables to investigate was inevitably forced and certainly not exhaustive; there are obviously other important executive functions that have to be examined thoroughly and many more issues that need to be explored in detail to unveil their cognitive and neural architecture. Increasing efforts are being made to achieve this goal, and the latent variable approach used here will certainly be very fruitful in this challenge.

Acknowledgments

The authors thank Mariagrazia Capizzi for her assistance in data collection.

Funding

This work was supported by the (FP7/2007-2013) European Research Council Starting Grant LEX-MEA [grant number 313692] to A. V.

References

Ambrosini, E., & Vallesi, A. (2016). Asymmetry in prefrontal resting-state EEG spectral power underlies individual differences in phasic and sustained cognitive control. *NeuroImage*, *124,* 843–857. http://doi.org/10.1016/j.neuroimage.2015.09.035

Ambrosini, E., & Vallesi, A. (2017). Domain-general Stroop Performance and Hemispheric Asymmetries: A Resting-state EEG Study. *Journal of Cognitive Neuroscience*, *29(5),* 769–779. http://doi.org/10.1162/jocn_a_01076

Arbula, S., Pacella, V., De Pellegrin, S., Rossetto, M., Denaro, L., D'Avella, D., Della Puppa, A., Vallesi, A. (2017). Addressing the selective role of distinct prefrontal areas in response suppression: A study with brain tumor patients. *Neuropsychologia*, *100,* 120-130. http://doi.org/10.1016/j.neuropsychologia.2017.04.018

Asbjørnsen, A. E., & Bryden, M. . (1998). Auditory attentional shifts in reading-disabled students: quantification of attentional effectiveness by the Attentional Shift Index. *Neuropsychologia, 36(2),* 143–148. http://doi.org/10.1016/S0028-3932(97)00090-0

Azur, M. J., Stuart, E. A., Frangakis, C., Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res, 20(1), 40-49. http://doi.org/10.1002/mpr.329

Babcock, L., & Vallesi, A. (2015). Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Bilingualism: Language and Cognition*, *20(2),* 403–417. http://doi.org/10.1017/S1366728915000735

Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology Section A*, *49(1),* 5–28. http://doi.org/10.1080/713755608

Badre, D., & Wagner, A.D., (2006). Computational and neurobiological mechanisms underlying cognitive flexibility. *Proc. Natl. Acad. Sci. U. S. A.* 7186–7191. https://doi.org/10.1073/pnas.0509550103

Banich, M. T. (2009). Executive Function. *Current Directions in Psychological Science, 18(2),* 89–94. http://doi.org/10.1111/j.1467-8721.2009.01615.x

Barkley, R. A. (2012). Executive functions : what they are, how they work, and why they evolved. New York, NY: Guilford Press.

Bettcher, B. M., Mungas, D., Patel, N., Elofson, J., Dutt, S., Wynn, M., Watson, C. L., Stephens, M., Walsh, C. M., & Kramer, J. H. (2016). Neuroanatomical substrates of executive functions:

beyond prefrontal structures. *Neuropsychologia*, *85,* 100-109. https://doi.org/10.1016/j.neuropsychologia.2016.03.001

Braver, T. S., Reynolds, J. R., & Donaldson, D. I., (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, *39,* 713–726. https://doi.org/10.1016/S0896-6273(03)00466-5

Burgess, P. W. (1997). Theory and Methodology in Executive Function Research. P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). East Sussex, UK: Psychology Press Ltd.

Burgess, P. W., & Shallice, T. (1996). Response suppression, Initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34(4),* 263–272. http://doi.org/10.1016/0028-3932(95)00104-2

Buuren, S. V., & Groothuis-Oudshoorn, K. (2010). Mice: Multivariate imputation by chained equations in R. *Journal of statistical software*, 1-68. http://doi.org/10.18637/jss.v045.i03

Capizzi, M., Ambrosini, E., Arbula, S., Mazzonetto, I., & Vallesi, A. (2016a). Electrophysiological Evidence for Domain-General Processes in Task-Switching. *Frontiers in Human Neuroscience*, *10,* 124. http://doi.org/10.3389/fnhum.2016.00124

Capizzi, M., Ambrosini, E., Arbula, S., Mazzonetto, I., & Vallesi, A. (2016b). Testing the domain-general nature of monitoring in the spatial and verbal cognitive domains. *Neuropsychologia*, *89,* 83–95. http://doi.org/10.1016/j.neuropsychologia.2016.05.032

Capizzi, M., Fehér, K., Penolazzi, B., & Vallesi, A. (2015). Task-switching preparation across semantic and spatial domains: An event-related potential study. *Biological Psychology*, *110,* 148–158. http://doi.org/10.1016/j.biopsycho.2015.06.011

Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23(2),* 201–216. http://doi.org/10.1016/j.acn.2007.08.010

Cieslik, E. C., Mueller, V. I., Eickhoff, C. R., Langner, R., & Eickhoff, S. B. (2015). Three key regions for supervisory attentional control: evidence from neuroimaging meta-analyses. *Neuroscience & Biobehavioral Reviews*, *48,* 22-34. https://doi.org/10.1016/j.neubiorev.2014.11.003

Cook, R. D., & Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman and Hall. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/37076.

Dang, C.-P., Braeken, J., Colom, R., Ferrer, E., & Liu, C. (2014). Why is working memory related to intelligence? Different contributions from storage and processing. *Memory*, *22(4),* 426–441. http://doi.org/10.1080/09658211.2013.797471

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11(3),* 201. http://doi.org/10.1038/nrn2793

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64(1),* 135–168. http://doi.org/10.1146/annurev-psych-113011-143750

Drazin, D. H. (1961). Effects of foreperiod, foreperiod variability, and probability of stimulus occurrence on simple reaction time. *Journal of Experimental Psychology*, *62,* 43–50. http://dx.doi.org/10.1037/ h0046860

Dreher, J. C., Koechlin, E., Ali, S. O., Grafman, J., (2002). The roles of timing and task order during task switching. *NeuroImage*, *17,* 95–109. https://doi.org/10.1006/nimg.2002.1169

Eslinger, P. J. (1996). Conceptualizing, describing, and measuring components of executive function: A summary. In G. R. Lyon & N. A. Krasnegor (Eds.), Attention, memory, and executive function (pp. 367-395). Baltimore, MD, US: Paul H Brookes Publishing Co. http://psycnet.apa.org/psycinfo/1995-98902-019

Fillmore, C. J., Kempler, D., & Wang, W. S.-Y. (1979). Individual differences in language ability and language behavior. New York, NY: Academic Press.

Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of Prefrontal and Anterior Cingulate Regions in Decision-Making Processes Shared by Memory and Nonmemory Tasks. *Cerebral Cortex*, *16(11),* 1623–1630. http://doi.org/10.1093/cercor/bhj097

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). “Sculpting the Response Space”—An Account of Left Prefrontal Activation at Encoding. *NeuroImage*, *12(4),* 404–417. http://doi.org/10.1006/nimg.2000.0633

Floden, D., Alexander, M. P., Kubu, C. S., Katz, D., & Stuss, D. T. (2008). Impulsivity and risk-taking behavior in focal frontal lobe lesions. *Neuropsychologia*, *46(1),* 213–23. http://doi.org/10.1016/j.neuropsychologia.2007.07.020

Fournier-Vicente, S., Larigauderie, P., & Gaonac’h, D. (2008). More dissociations and interactions within central executive functioning: A comprehensive latent-variable analysis. *Acta Psychologica*, *129(1),* 32–48. http://doi.org/10.1016/j.actpsy.2008.04.004

Friedman, N. P., & Miyake, A. (2004). The Relations Among Inhibition and Interference Control Functions: A Latent-Variable Analysis. *Journal of Experimental Psychology: General*, *133(1),* 101–135. http://doi.org/10.1037/0096-3445.133.1.101

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not All Executive Functions Are Related to Intelligence. *Psychological Science*, *17(2),* 172–179. http://doi.org/10.1111/j.1467-9280.2006.01681.x

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, *43(2),* 226–36. http://doi.org/10.3758/s13421-014-0461-7

Gold, B. T., Powell, D. K., Xuan, L., Jicha, G. A., & Smith, C. D. (2010). Age-related slowing of task switching is associated with decreased integrity of frontoparietal white matter. *Neurobiology of Aging*, *31(3),* 512–22. http://doi.org/10.1016/j.neurobiolaging.2008.04.005

Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annu Rev Psychol, 60,* 549–576. http://doi.org/10.1146/annurev.psych.58.110405.085530

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives*. Structural Equation Modeling: A Multidisciplinary Journal*, *6(1),* 1–55. http://doi.org/10.1080/10705519909540118

Hugdhal, K., Westerhausen, R., Alho, K., Medvedev, S., Laine, M., & Hämäläinen, H. (2009). Attention and cognitive control: Unfolding the dichotic listening story. *Scandinavian Journal of Psychology*, *50(1),* 11–22. http://doi.org/10.1111/j.1467-9450.2008.00676.x

Jurado, M. B., & Rosselli, M. (2007). The Elusive Nature of Executive Functions: A Review of our Current Understanding. *Neuropsychology Review*, *17(3),* 213–233. http://doi.org/10.1007/s11065-007-9040-z

Kim, C., Johnson, N. F., Cilles, S. E., Gold, B. T., (2011a). Common and distinct mechanisms of cognitive flexibility in prefrontal cortex. *J Neurosci, 31,* 4771–4779. https://doi.org/10.1523/JNEUROSCI.5923-10.2011

Kim, M. J., Gee, D. G., Loucks, R. A., Davis, F. C., & Whalen, P. J. (2011b). Anxiety dissociates dorsal and ventral medial prefrontal cortex functional connectivity with the amygdala at rest. Cerebral Cortex, *21(7),* 1667–73. http://doi.org/10.1093/cercor/bhq237

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, *302(5648),* 1181–1185. http://doi.org/10.1126/science.1088545

Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., Turkeltaub, P.E., Kochunov, P., Fox, P. T. (2005). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping*, *25,* 155–164. http://doi.org/10.1002/hbm.20136

Laird, A. R., McMillan, K. M., Lancaster, J. L., Kochunov, P., Turkeltaub, P. E., Pardo, J. V., Fox, P. T. (2005). A comparison of label-based review and ALE meta-analysis in the Stroop task. *Human Brain Mapping*, *25,* 6–21. https://doi.org/10.1002/hbm.20129

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science, *288(5472),* 1835–8. https://doi.org/ 10.1126/science.288.5472.1835

MacPherson, S., Gillebert, C., Robinson, G., Vallesi, A. (2019). Intra- and Inter-individual Variability of Executive Functions: Determinant and Modulating Factors in Healthy and Pathological Conditions. *Front Psychol, 10,* 432. https://doi.org/10.3389/fpsyg.2019.00432

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, *95(4),* 833. http://dx.doi.org/10.1037/0022-0663.95.4.833

McCloskey, G., & Perkins, L. A. (2013). *Essentials of executive function assessment*. New York, NY: John Wiley & Sons.

Meiran, N. (2014). The task-cuing paradigm: A user's guide. In Grange J.A., Houghton G. (Eds.), *Task switching and cognitive control* (pp. 45–73). New York: Oxford University Press.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24(1),* 167–202. http://doi.org/10.1146/annurev.neuro.24.1.167

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science, 21(1),* 8–14. http://doi.org/10.1177/0963721411429458

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. http://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130(4),* 621–640. http://doi.org/10.1037/0096-3445.130.4.621

Mostofsky, S. H., Simmonds, D. J. (2008). Response inhibition and response selection: two sides of the same coin. *J Cogn Neurosci, 20,* 751–761. https://doi.org/10.1162/jocn.2008.20500

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, *5(2),* 159-168. https://doi.org/10.1207/s15327574ijt0502_4

Naber, M., Vedder, A., Brown, S. B. R. E., & Nieuwenhuis, S. (2016). Speed and Lateral Inhibition of Stimulus Processing Contribute to Individual Differences in Stroop-Task Performance. *Front Psychol, 7,* 822. https://doi.org/10.3389/fpsyg.2016.00822

Neumann, J., Lohmann, G., Derrfuss, J., & von Cramon, D. Y. (2005). Meta-analysis of functional imaging data using replicator dynamics. *Human Brain Mapping*, *25,* 165–173. https://doi.org/10.1002/hbm.20133

Neumann, J., von Cramon, D. Y., & Lohmann, G. (2008). Model-based clustering of meta-analytic functional imaging data. *Human Brain Mapping*, *29,* 177–192. https://doi.org/10.1002/hbm.20380

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126(2),* 220–246. http://doi.org/10.1037/0033-2909.126.2.220

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9(1),* 97–113. http://doi.org/10.1016/0028-3932(71)90067-4

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360(1456),* 781–95. http://doi.org/10.1098/rstb.2005.1631

Phillips, L. H. (1997). Do "frontal tests" measure executive function? Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 191–213). Hove, UK: Psychology Press.

Puccioni, O., & Vallesi, A. (2012a). Conflict resolution and adaptation in normal aging: The role of verbal intelligence and cognitive reserve. *Psychology and Aging*, *27(4),* 1018–1026. http://doi.org/10.1037/a0029106

Puccioni, O., & Vallesi, A. (2012b). High cognitive reserve is associated with a reduced age-related deficit in spatial conflict resolution. *Frontiers in Human Neuroscience*, *6,* 327. http://doi.org/10.3389/fnhum.2012.00327

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rau, H. K., Suchy, Y., Butner, J. E., & Williams, P. G. (2016). Latent profiles of executive functioning in healthy young adults: evidence of individual differences in hemispheric asymmetry. *Psychological Research*, *80(6),* 997-1019. https://doi.org/10.1007/s00426-015-0706-5

Robinson, G. A., Cipolotti, L., Walker, D. G., Biggs, V., Bozzali, M., & Shallice, T. (2015). Verbal suppression and strategy use: a role for the right lateral prefrontal cortex? *Brain*, *138(4),* 1084–1096. http://doi.org/10.1093/brain/awv003

Rosseel, Y. (2012). lavaan : An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48(2),* 1–36. http://doi.org/10.18637/jss.v048.i02

Rousseeuw, P. J., & Verboven, S. (2002). Robust estimation in very small samples. *Computational Statistics & Data Analysis*, *40(4),* 741–758. http://doi.org/10.1016/S0167-9473(02)00078-6

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive Functioning as a Potential Mediator of Age-Related Cognitive Decline in Normal Adults. *Journal of Experimental Psychology: General*, *132(4),* 566–594. http://doi.org/10.1037/0096-3445.132.4.566

Shallice, T., Burgess, P., & Robertson, I. (1996). The Domain of Supervisory Processes and Temporal Organization of Behaviour [and Discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *351(1346),* 1405–1412. http://doi.org/10.1098/rstb.1996.0124

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994). Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature*, *368(6472),* 633–635. http://doi.org/10.1038/368633a0

Shallice, T., Stuss, D. T., Picton, T. W., Alexander, M. P., & Gillingham, S. (2007). Multiple effects of prefrontal lesions on task-switching. *Frontiers in Human Neuroscience*, *1,* 2. http://doi.org/10.3389/neuro.09.002.2007

Shallice, T., Stuss, D. T., Picton, T. W., Alexander, M. P., & Gillingham, S. (2008). Mapping task switching in frontal cortex through neuropsychological group studies. *Frontiers in Neuroscience*, *2(1),* 79–85. http://doi.org/10.3389/neuro.01.013.2008

Steinborn, M. B., & Langner, R. (2012). Arousal modulates temporal preparation under increased time uncertainty: Evidence from higher order sequential foreperiod effects. *Acta Psychologica*, *139,* 65–76. http://dx.doi.org/10.1016/j.actpsy.2011.10.010

Stuss, D. T. (2011). Functions of the Frontal Lobes: Relation to Executive Functions. *Journal of the International Neuropsychological Society*, *17(5),* 759–765. http://doi.org/10.1017/S1355617711000695

Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive syndrome ? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362(1481),* 901–915. http://doi.org/10.1098/rstb.2007.2096

Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., Borowiec, A., Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, *43(3),* 396–417. http://doi.org/10.1016/j.neuropsychologia.2004.06.010

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of the New York Academy of Sciences*, *769,* 191–212. http://doi.org/10.1111/j.1749-6632.1995.tb38140.x

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics. Boston: Allyn and Bacon.

Tarantino, V., Mazzonetto, I., Formica, S., Causin, F., & Vallesi, A. (2017). The Neural Bases of Event Monitoring across Domains: a Simultaneous ERP-fMRI Study. *Frontiers in Human Neuroscience*, *11,* 376. https://doi.org/10.3389/fnhum.2017.00376

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the*

National Academy of Sciences of the United States of America, 94(26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792

Triviño, M., Correa, A., Arnedo, M., & Lupiañez, J. (2010). Temporal orienting deficit after prefrontal damage. *Brain*, *133,* 1173-1185. https://doi.org/10.1093/brain/awp346

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37(3),* 498–505. http://doi.org/10.3758/BF03192720

Unsworth, N., Miller, J. D., Lakey, C. E., Young, D. L., Meeks, J. T., Campbell, W. K., & Goodie, A. S. (2009). Exploring the Relations Among Executive Functions, Fluid Intelligence, and Personality. *Journal of Individual Differences*, *30(4),* 194–200. http://doi.org/10.1027/1614-0001.30.4.194

Vallesi, A. (2012). Organisation of executive functions : Hemispheric asymmetries. *Journal of Cognitive Psychology*, *24(4),* 367–386. http://doi.org/10.1080/20445911.2012.678992

Vallesi, A., Arbula, S., & Bernardis, P. (2014). Functional dissociations in temporal preparation: Evidence from dual-task performance. *Cognition*, *130(2),* 141–151. http://doi.org/10.1016/j.cognition.2013.10.006

Vallesi, A., Arbula, S., Capizzi, M., Causin, F., & D'Avella, D. (2015). Domain-independent neural underpinning of task-switching: An fMRI investigation. *Cortex*, *65,* 173–183. https://doi.org/10.1016/j.cortex.2015.01.016

Vallesi A., Lozano V., Correa A. (2013). Dissociating temporal preparation processes as a function of the inter-trial interval duration. *Cognition*, *127,* 22–30. https://doi.org/10.1016/j.cognition.2012.11.011

Vallesi A., Mastrorilli E., Causin F., D'Avella D., Bertoldo A. (2016) White matter and task-switching in young adults: A Diffusion Tensor Imaging study. *Neuroscience, 329,* 349-362. https://doi.org/10.1016/j.neuroscience.2016.05.026

Vallesi, A., McIntosh, A. R., Alexander, M., Stuss, D. T. (2009). FMRI evidence of a functional network setting the criteria for withholding a response. *Neuroimage*, *45,* 537-548. https://doi.org/10.1016/j.neuroimage.2008.12.032

Vallesi, A., McIntosh, A. R., Shallice, T., & Stuss, D. T. (2009). When time shapes behavior: fMRI evidence of brain correlates of temporal monitoring. *J Cogn Neurosci, 21,* 1116-1126. https://doi.org/10.1162/jocn.2009.21098

Vallesi, A., McIntosh, A. R., Crescentini, C., & Stuss, D. T. (2012). fMRI investigation of speed-accuracy strategy switching. *Human Brain Mapping*, *33(7),* 1677–1688. http://doi.org/10.1002/hbm.21312

Vallesi, A., Mussoni, A., Mondani, M., Budai, R., Skrap, M., & Shallice, T. (2007a). The neural basis of temporal preparation: Insights from brain tumor patients. *Neuropsychologia, 45(12),* 2755–2763. http://doi.org/10.1016/j.neuropsychologia.2007.04.017

Vallesi A., Shallice T. (2007). Developmental dissociations of preparation: deconstructing variable foreperiod phenomena. *J Exp Psychol Hum Percept Perform, 33(6),* 1377–1388. http://doi.org/10.1037/0096-1523.33.6.1377

Vallesi, A., Shallice, T., & Walsh, V. (2007b). Role of the Prefrontal Cortex in the Foreperiod Effect: TMS Evidence for Dual Mechanisms in Temporal Preparation. *Cerebral Cortex*, *17(2),* 466–474. http://doi.org/10.1093/cercor/bhj163

Verboven, S., & Hubert, M. (2005). LIBRA: a MATLAB library for robust analysis. *Chemometrics and Intelligent Laboratory Systems*, *75(2),* 127–136. http://doi.org/10.1016/j.chemolab.2004.06.003

Verboven, S., & Hubert, M. (2010). MATLAB library LIBRA. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2(4),* 509–515. http://doi.org/10.1002/wics.96

Wang, Y., Kuhl, P.K., Chen, C., & Dong, Q., (2009). Sustained and transient language control in the bilingual brain. *NeuroImage*, *47,* 414–422. https://doi.org/10.1016/j.neuroimage.2008.12.055

Weiskrantz, L., (1992). Introduction: Dissociated Issues. In A. D. Milner & M. D. Rugg (Eds*.), The neuropsychology of consciousness* (pp. 1-10). London: Academic Press.

Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). The relation between a multicomponent working memory and intelligence: The roles of central executive and short-term storage functions. *Intelligence*, *53,* 166–180. http://doi.org/10.1016/j.intell.2015.10.007

Tables

Table 1. Zero-order correlations among executive function scores

		1	2	3	4	5	6	7	8	9	10	11	12
1	ver_Swi		**<.001**	**.014**	.649	.109	**.041**	.194	.124	.062	.579	.491	.549
2	spa_Swi	**.411**		**<.001**	**.009**	**.023**	**.043**	.535	**.041**	.647	.342	.344	.193
3	c-s_Swi	**.197**	**.350**		.347	**.024**	.433	.861	.119	.643	.067	**.015**	.654
4	ver_Stroop	.037	**.206**	.076		**.015**	.246	.967	.840	.110	**.003**	.787	.880
5	spa_Stroop	.128	**.182**	**.180**	**.194**		.631	.093	.187	.312	.090	.080	.336
6	DL_zASI	**.164**	**.162**	.063	-.093	-.039		.510	.391	.979	.810	.876	.943
7	ver_Mon	.104	-.050	-.014	.003	.134	-.053		**.006**	**.002**	**.001**	**.001**	.244
8	spa_Mon	.123	**.163**	-.125	.016	.106	.069	**.220**		.051	**.004**	**.002**	.111
9	ver_Mix	-.149	.037	-.037	.128	.081	-.002	**.250**	.156		**<.001**	.062	.295
10	spa_Mix	.045	-.076	-.147	**.237**	.136	.019	**.261**	**.228**	**.314**		**.002**	.626
11	c-s_Mix	.055	.076	**-.194**	.022	.140	-.013	**.262**	**.251**	.149	**.244**		**.023**
12	Foreperiod	.048	.104	.036	-.012	.077	.006	.094	.128	.084	-.039	**.182**	

Notes: Values of the Pearson's correlation coefficients and the corresponding *p*-values are shown in the lower and upper triangles, respectively. Significant correlations are indicated in bold. Ver_, verbal task; spa_, spatial task; c-s_, color-shape task; Swi, switch cost; Stroop, Stroop effect; DL_zASI, dichotic listening task, standardized attentional shift index; Mon, monitoring cost; Mix, mixing cost; Foreperiod, foreperiod task (see the section "Behavioral Tasks, Procedures and Data Preparation" for details).

Table 2. Fit statistics for the confirmatory factor analysis models discussed

Model[a]	df	χ^2	*P*	AIC	CFI	RMSEA	SRMR
CS_M×L_R[b]	45	54.15	.165	5237	.940	.036	0.064
L_R[b]	50	104.20	< .001	5277	.643	.083	0.099
CS_M (Figure 2)	50	60.77	.142	5234	.929	.037	0.067
Unrelated CS_M	51	62.67	.127	5233	.923	.038	0.072
Unitary CS_M[c]	51	97.11	<.001	5268	.696	.076	0.090

Notes: Non-significant χ^2 statistics indicate reasonable fits to the data. The smaller the Akaike Information Criterion (AIC) value is, the better the model fit is. Higher values of the Bentler's comparative fit index (CFI) indicate better fit, with CFI > .90 indicating an acceptable fit. Lower values of root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) indicate better fit, with SRMR < .08 and RMSEA < .06 indicating an acceptable fit to the data.

[a] The left indentation indicates nested models. L_R, Left- vs. Right-lateralized domains model; CS_M, Criterion Setting vs. Monitoring model; CS_M×L_R, interaction model.

[b] The fit of these models was problematic as the covariance matrix was not positive definite.

[c] Note that this model is equivalent to the L=R model (see Domain-based models).

Figure captions

Figure 1. Experimental tasks.

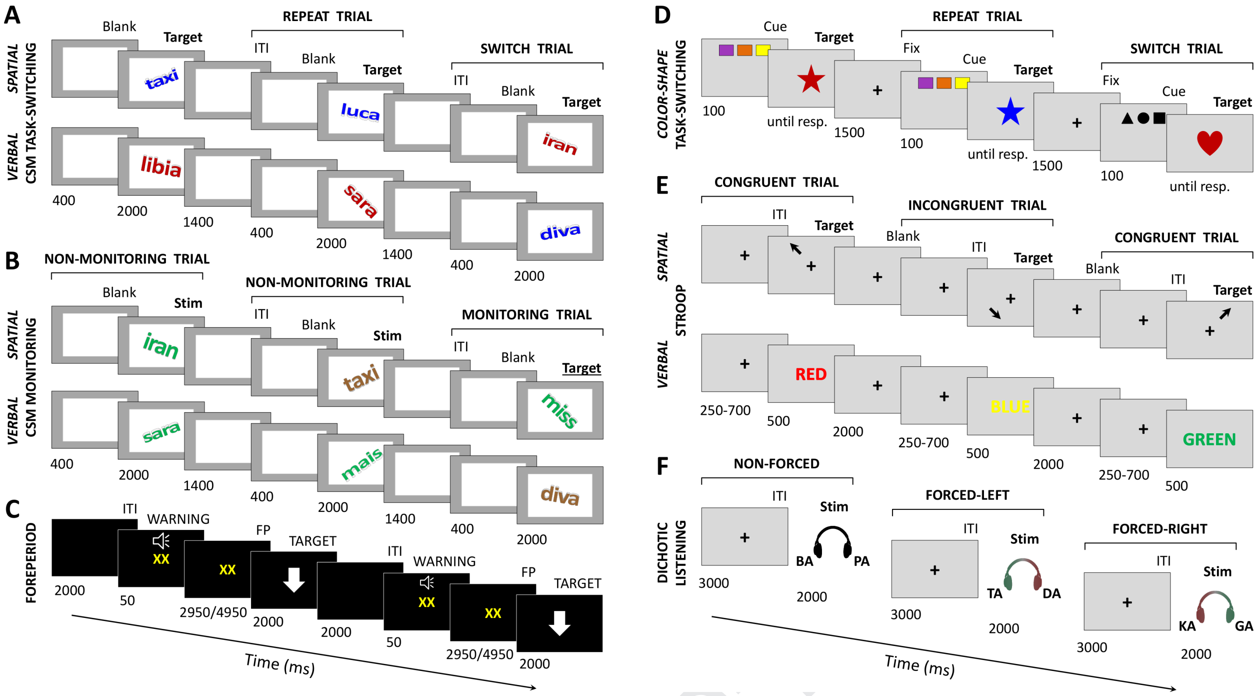
The figure shows the trial structure of the experimental tasks we used. A) Verbal (lower row) and spatial (upper row) task-switching versions of the Criterion-setting/monitoring (CSM) paradigm. The first three trials in a mixed block are shown, including a repeat and a switch trial. B) Verbal (lower row) and spatial (upper row) monitoring tasks of the CSM paradigm. Two non-monitoring and a monitoring trials are shown. C) Foreperiod task. D) Color-shape task-switching paradigm. The first three trials in a mixed block are shown, including a repeat and a switch trial. E) Verbal (lower row) and spatial (upper row) Stroop tasks. Two congruent trials and an incongruent trial are shown. F) Dichotic listening task. Trials for the non-forced, forced-left, and forced-right conditions are shown. ITI, inter-trial interval; Fix, fixation; Stim, stimuli; FP, foreperiod. See "Behavioral Tasks" section for more details.

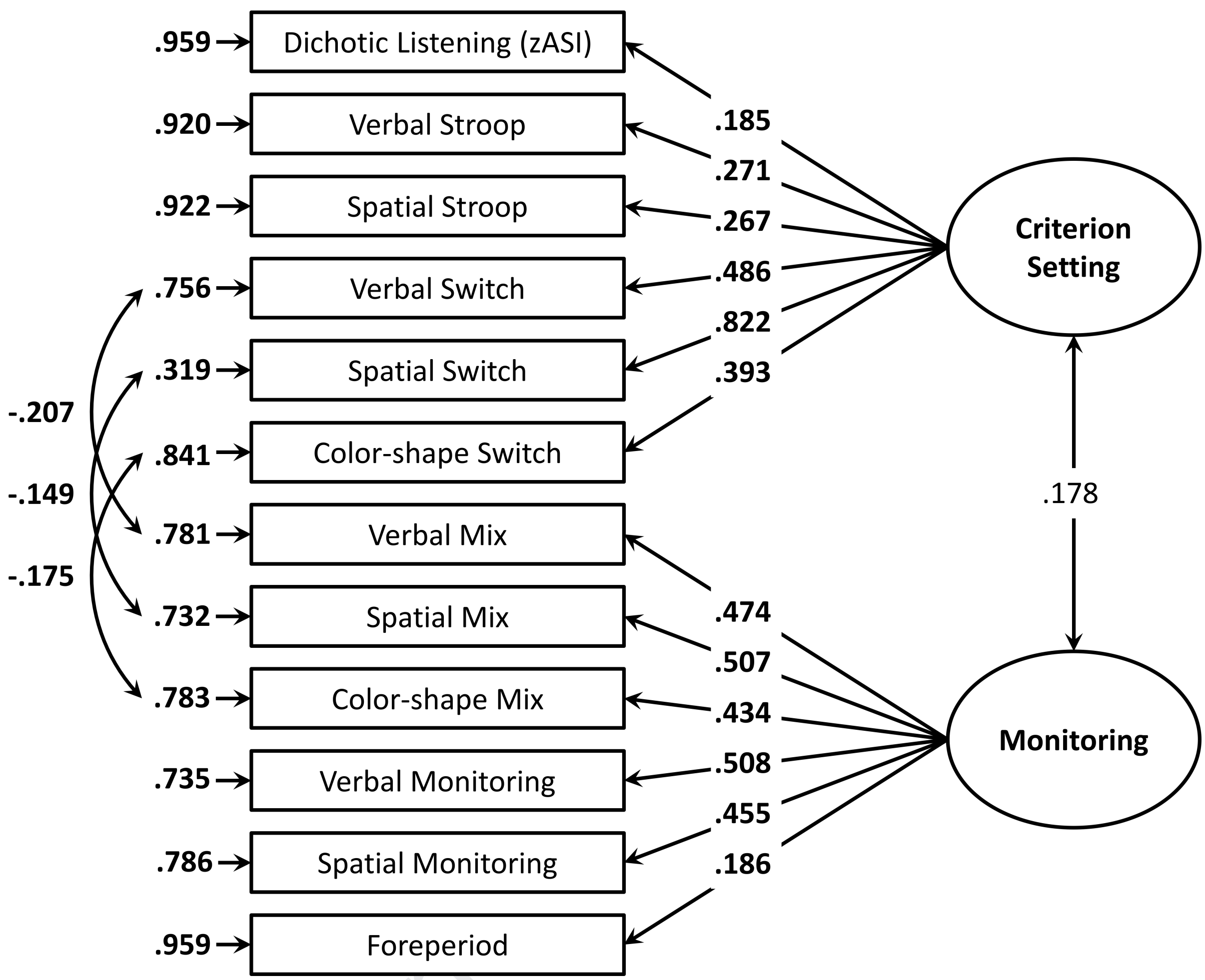
Figure 2. The two-factor process-based Criterion-Setting-Monitoring (CS_M) model.

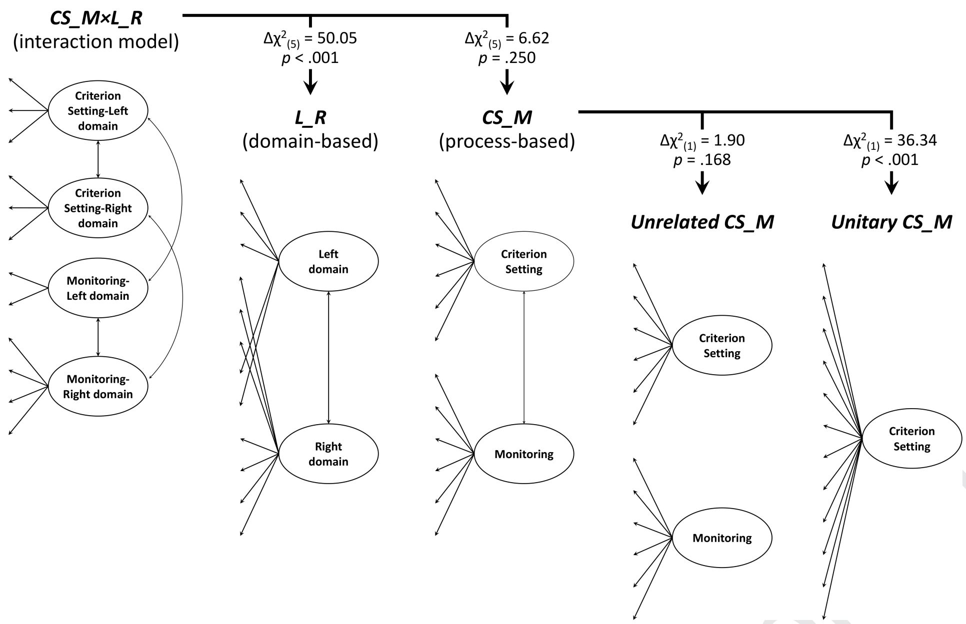
The rectangles represent the observed executive function scores, while the ovals represent the latent variables. The numbers over the straight, single-headed arrows that go from the latent variables to the executive function scores are the standardized factor loadings. The numbers at the end of the smaller arrows on the left are the error variances for each executive function score due to measurement error and task-specific requirements. The numbers over the curve, double-headed arrows on the left indicate the correlations of the task error variances. The number next to the double-headed arrow connecting the latent variables shows the correlation between them. Significant factor loadings and correlations are indicated in bold.

Figure 3. Representation and results of the comparisons between the tested models.

Schematic representations of the different models we tested and their relations. Nested models are indicated by thick arrows and the corresponding statistics are shown for difference tests on their fit; significant results indicate that the fuller model (on the left) has a significantly better fit than the nested one (on the right). L_R, Left- vs. Right-lateralized domains model; CS_M, Criterion Setting vs. Monitoring model; CS_M×L_R, process by domain interaction model (see text for more details).




.959
Dichotic Listening (zASI)
.920
Verbal Stroop
.922
Spatial Stroop
.756
Verbal Switch
.319
Spatial Switch
.841
Color-shape Switch
.781
Verbal Mix
.732
Spatial Mix
.783
Color-shape Mix
.735
Verbal Monitoring
.786
Spatial Monitoring
.959
Foreperiod
-.207
-.149
-.175
.185
.271
.267
.486
.822
.393
.474
.507
.434
.508
.455
.186
Criterion Setting
.178
Monitoring


CS_M×L_R
(interaction model)
Δχ²(5) = 50.05
p < .001
Δχ²(5) = 6.62
p = .250
L_R
(domain-based)
CS_M
(process-based)
Δχ²(1) = 1.90
p = .168
Δχ²(1) = 36.34
p < .001
Unrelated CS_M
Unitary CS_M
Criterion Setting-Left domain
Criterion Setting-Right domain
Monitoring-Left domain
Monitoring-Right domain
Left domain
Right domain
Criterion Setting
Monitoring
Criterion Setting
Monitoring
Criterion Setting

CRediT Author Statement

Ettore Ambrosini: Conceptualization, Methodology, Software, Formal Analysis, Visualization, Writing – Original Draft; Data Curation; **Sandra Arbula:** Conceptualization, Methodology, Software, Investigation, Writing – Review & Editing; **Chiara Rossato:** Investigation, Writing – Review & Editing; **Valentina Pacella:** Investigation, Writing – Review & Editing; **Antonino Vallesi:** Conceptualization, Methodology, Software, Writing – Review & Editing, Supervision, Project Administration, Funding Acquisition