



**HAL**  
open science

# Cognitive Flexibility Or Flexibilities? Insights From A Classroom Study

Laure Baudier, Evelyne Clément, Emmanuel Sander, Hippolyte Gros

► **To cite this version:**

Laure Baudier, Evelyne Clément, Emmanuel Sander, Hippolyte Gros. Cognitive Flexibility Or Flexibilities? Insights From A Classroom Study. 2024. hal-04586020

**HAL Id: hal-04586020**

**<https://hal.science/hal-04586020>**

Preprint submitted on 24 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

# Cognitive Flexibility Or Flexibilities? Insights From A Classroom Study.

Laure Baudier, Evelyne Clément, Emmanuel Sander, and Hippolyte Gros.

## Abstract

There is an ongoing debate in the scientific community regarding the nature of cognitive flexibility. It is either seen as a general executive process (*set shifting*), as a dimensional construct composed of reactive flexibility and spontaneous flexibility, or as a task-dependent skill. To help decide between these views, we analyzed the performances of 86 first-graders across four Cognitive Flexibility tasks. Two tasks focused on reactive flexibility (*rule-shifting* and *predicate-shifting*), and two tasks targeted spontaneous flexibility (*role-shifting* and *divergent thinking*). We also assessed children's performance on three Executive Functions tasks (*cognitive inhibition*, *verbal working memory*, and *visuomotor processing speed*) to investigate the extent to which they correlate with cognitive flexibility. Significant — though moderate — positive correlations emerged among five out of the six pairs of cognitive flexibility tasks, regardless of whether they involved a reactive or spontaneous use of cognitive flexibility. Besides, none of the executive functions tasks systematically correlated with the four measures of cognitive flexibility. On the other hand, Confirmatory Factorial Analyses did not state whether cognitive flexibility is a unifactorial or bi-factorial construct. Overall, our results suggest that cognitive flexibility performance may reflect the recruitment of several task-dependent transversal processes. Implications with regard to the three postulates as well as future directions are discussed.

**Keywords:** Cognitive flexibility; Reactive flexibility; Spontaneous flexibility; Correlations; Confirmatory Factorial Analyses; Education; Executive control; First-graders

## Introduction

Does the ability to deal with an unexpected event in a given situation guarantee an individual's overall capacity to face any contextual change? In our constantly changing environment, adapting behaviors to varying external demands is an everyday challenge. By directing attention to relevant contextual information, cognitive flexibility makes it possible to change one's way of thinking, to adapt to new circumstance, and to modify representations or strategies according to different goals (e.g., Chevalier et al., 2012; Clément, 2009; Deák, 2003; Diamond, 2013; Ionescu, 2012, 2017; Miyake et al., 2000; Zelazo et al., 2003). For instance, adapting to a change of instructions while performing a task, grasping an unexpected opportunity, or coming up with an original idea, are adaptive behaviors that all rely on cognitive flexibility. Cognitive flexibility is so central that, together with inhibitory control and working memory, it predicts several factors of life-happiness, such as achievement, health, wealth, success, and quality of life (for a review see Diamond & Ling, 2016).

Cognitive flexibility has been shown to support the development of several skills from a young age, such as creative thinking (e.g., Arán Filippetti & Krumm, 2020; Ebersbach & Hagedorn, 2011; Pan & Yu, 2018; Runco, 2004), lexical and conceptual knowledge (e.g., Blaye

& Bonthoux, 2001; Blaye & Maintenant, 2008; Deák, 2003; Duvignau et al., 2007), problem solving (Clément, 2009, 2022; Gros, Thibaut, & Sander, 2020; Raynal, 2022; Zelazo & Frye, 1998), as well as many aspects of school achievement such as mathematics, reading, and literacy (e.g., Arán Filippetti & Richaud, 2017; Bull & Scerif, 2001; Gros & Gvozdic, 2022; Iacono, Gros, & Clément, 2022; Titz & Karbach, 2014; Yeniad et al., 2013). For instance, it has been suggested that cognitive flexibility development contributes to reading fluency in early readers by helping them to switch between phonological and semantic word properties (Cartwright et al., 2019).

Given the need to develop flexible thinking from an early age, many attempts are made in translational research to promote its expression in an educational setting. However, the interventions often target vastly different abilities depending on the research framework they fall within. Indeed, there is currently a heated debate in the scientific community regarding *what exactly* characterizes a flexible cognitive system, without a consensual definition emerging from the wealth of research focusing on cognitive flexibility (Hohl & Dolcos, 2024; Ionescu, 2012). Instead, three main lines of studies emerge from the literature, each yielding specific conclusions about the nature of this construct, resulting in different experimental designs and training interventions. In practical terms, this raises the question: how should cognitive flexibility be trained? Should it be conceived of as (i) A *general executive process*, whose improvement in a given domain would potentially benefit every other context (for a review of studies carried out under this approach, see Diamond & Ling, 2016)? (ii) A *dimensional construct* encompassing two distinct types of flexible behaviors (whether reacting to an explicit environmental constraint, or spontaneously initiating change – as recently suggested by Arán Filippetti & Krumm, 2020)? or (iii) A *task-dependent skill* whose improvement may only lead to narrow transfer, benefitting tasks carried out under a similar context (e.g., training cognitive flexibility in arithmetic word problem solving, as in Iacono et al., 2022; and Scheibling-Sève et al., 2022)? Considering the lack of a scientific consensus between these three perspectives, it seems important to reach a unified definition of cognitive flexibility before implementing educational interventions targeting this construct. In this perspective, we propose to review the arguments in favor of each of these accounts of cognitive flexibility.

## **A General Executive Process?**

According to an influential line of research in the field of cognitive control and executive functions, cognitive flexibility is a generalized cognitive process that emerges early in life and continues to develop during childhood and adolescence, following cortical development. Under this view, cognitive flexibility underpins flexible behaviors across a wide range of situations, and the individuals become increasingly flexible as their brain matures over age (e.g., Diamond, 2013; for a review see Ionescu, 2012).

This conception is grounded in pioneering neuropsychological works on cognitive control (Miyake et al., 2000). Using *executive tasks*, research in frontal lobe patients has shown that the ability to pursue changing or new task-goal relies on specific cortical and subcortical neural networks (Drewe, 1974; Milner, 1963; Nelson, 1976). Following these seminal works, current studies often consider that humans regulate their thoughts and behaviors through three core

executive functions: cognitive flexibility (defined as the shift of attention between relevant pieces of information), inhibitory control (the suppression of irrelevant information or responses) and working memory (information processing and updating) (e.g., Best, Miller, & Jones, 2009; Bull & Scerif, 2001; Espy et al., 2004; Lehto et al., 2003; Miyake et al., 2000). Under this framework, cognitive flexibility is thus conceived of as the ability to shift between tasks and/or stimuli, also called *set shifting* (Chevalier et al., 2010; Cragg & Chevalier, 2012; Diamond, 2013; Ravizza & Carter, 2008).

Several paradigms have thus been developed to assess the ability to shift from one task or rule to another (e.g., reversal of stimulus-response mapping, attentional shift between stimulus properties, Diamond 2013; Monsell, 2003; Zelazo et al., 2003). Depending on the paradigm, participants either need to switch back and forth between two tasks for each trial (*switching* tasks) or to shift to a new task after having performed the same task repeatedly in a series of trials (*rule-shifting* tasks). Children's cognitive flexibility is mostly assessed with *rule-shifting* tasks (e.g., Bunge & Zelazo, 2006; Davidson et al., 2006; Jacques et al., 1999; Kirkham, Cruess, & Diamond, 2003; Kloo et al., 2010; Perner & Lang, 2002). *Rule-shifting* tasks are considered an appropriate measure of children's cognitive flexibility because they involve less frequent shifts compared to *switching* paradigms, which are preferentially used with adults (Cepeda, Kramer, & Gonzalez de Sather, 2001; Davidson et al., 2006; Peng, Kirkham, & Mareschal, 2018). For instance, the most popular *rule-shifting* task among preschoolers is the Dimensional Change Card Sort (DCCS, Zelazo, 2006). The DCCS uses several cards depicting colored figures. Children are first asked to sort all the cards in two piles according to a first rule, before being asked to sort the cards again, according to a new rule. For example, they first are told to sort the red drawings on one side and the blue drawings on the other (regardless of their shape: truck or duck), then the truck drawings on one side and the duck drawings on the other (regardless of their color: blue or red). For school-aged children, cognitive flexibility is mostly assessed with the Wisconsin Card Sorting Test (WCST, Grant & Berg, 1948). The WCST is more complex than the DCCS, since its stimuli vary on a higher number of properties (Color, Shape, Number), thus inducing a higher number of rule changes.

A considerable amount of research using the DCCS and the WCST have studied the development of cognitive flexibility during childhood (Chelune & Baer, 1986; Frye, Zelazo, & Palfai, 1995; Kirkham, Cruess, & Diamond, 2003; Rosselli & Ardila, 1993; Shu et al., 2000; Zelazo et al., 2003). The steady development of cognitive flexibility over the years is showcased by children's ability to process an increasing number of sorting rules, as well as by decreasing perseverative errors (e.g., Davidson et al., 2006; Diamond, 2006, 2013; Zelazo & Frye, 1998; Zelazo et al., 2003). Developmental milestones have been identified using these two tasks. Regarding the DCCS, the ability to shift from the first to the second sorting rule is generally reached by 4- to 5-year-old children – while 3-year-old children persevere in categorizing according to the first rule even after having been explained the new rule (Frye, Zelazo, & Palfai, 1995; Zelazo, Frye, & Rapus, 1996; Zelazo et al., 2003). Similarly, results on the WCST indicate that children are able to shift between three rules and reach adult-like performance at around 10 years old (Chelune & Baer, 1986). This suggests that *set shifting* develops uniformly across individuals, which would support the hypothesis according to which cognitive flexibility is a generalized competence.

Increasing performance in *rule-shifting* tasks is often seen as closely related to the development of cognitive control, and more specifically to inhibitory control and working memory executive processes (Diamond, 2013; Marcovitch & Zelazo, 2009; Zelazo et al., 2003). Interestingly, the development of the prefrontal cortex has been shown to correlate with children's *rule-shifting* performance (e.g., Bunge & Zelazo, 2006; Ezeziel, Bosma, & Morton, 2013; Morton & Munakata, 2002; Zelazo & Frye, 1998). Besides, it would seem that there is a close relation between the improvement of executive functions and the development of the prefrontal cortex over age (for a review see Chevalier et al., 2010). This has been interpreted as evidence that children become more flexible because of the maturation of a generalized executive process within cognitive control (Cepeda, Cepeda, & Kramer, 2000; Davidson et al., 2006; Diamond, 2013; Muller et al., 2006; Zelazo et al., 2003).

It should be noted, however, that these conclusions stem from generalizations drawn from observations gathered using a narrow range of cognitive flexibility tasks. Indeed, many behaviors are deemed flexible (including adapting responses to a rule change, problem-solving, generating new ideas, word learning, Deák, 2003; Clément, 2009; Ionescu, 2012), but as noted by Clément (2022), *rule-shifting* and *switching* paradigms were designed with the aim to train and assess a general ability, not the full extent of what cognitive flexibility actually is. Implicitly, these tasks operate under the assumption that cognitive flexibility is a domain-general process relying on *set shifting* (the shift of attention between two tasks). Yet, this assumption is the subject of a heated debate in the literature regarding the nature of cognitive flexibility. Through the use of different flexibility tasks allowing for comparative analyses of individuals' performances across contexts, other studies have proposed alternative hypotheses regarding the nature of cognitive flexibility. Along this line, it is noteworthy that all of the above-mentioned studies used tasks in which cognitive flexibility was required to respond to changing tasks demands. Indeed, the ability to modify behavior following an external constraint has been described as a crucial aspect of cognitive flexibility. On the other hand, there appears to be another side of cognitive flexibility, involved in situations where no explicit indication for change is provided by the context. This aspect of cognitive flexibility, called *spontaneous flexibility*, is not under scrutiny in traditional cognitive flexibility tasks, and requires different paradigms to be identified.

## **Is Cognitive Flexibility A Dimensional Construct?**

Since pioneer neuropsychological works on dysexecutive disorders (Eslinger & Grattan, 1993), an important body of research has described cognitive flexibility as a dimensional construct with two subcomponents representing distinct skills: *reactive flexibility* and *spontaneous flexibility*. Reactive flexibility refers to the ability to adapt behaviors to environmental demands (e.g., an unforeseen event, a rule change). It is thought to rely on *set shifting* (Miyake et al., 2000). Spontaneous flexibility describes appropriate response changes initiated without any outside prompt (e.g., spontaneously considering a barrel as a musical instrument instead of a container, Pan & Yu, 2018). It requires divergent thought. Reactive flexibility is mostly assessed with the *switching* and *rule-shifting* tasks described above, while spontaneous flexibility is typically grasped with *divergent thinking* and *verbal fluency* paradigms (Clément, 2009, 2022; Eslinger & Grattan, 1993). *Divergent thinking* tasks require imagining a variety of

new uses for an object commonly used for a specific purpose (e.g., a brick, a pencil, a hanger, Gilhooly et al., 2007). *Verbal fluency* tasks assess the ability to propose as many words as possible beginning with the same letter (e.g., ‘P’, ‘V’), or belonging to the same category (e.g., Animals, Fruits) – letters and categories change across blocks (e.g., Cardebat et al., 1990).

In their seminal neuropsychological study in 1993, Eslinger and Grattan found evidence for this dichotomy by identifying neural networks specific to both components in adults with localized brain lesions. Depending on the lesions, deficits in reactive flexibility or spontaneous flexibility have been identified, each involving different cortical and subcortical networks (Eslinger & Grattan, 1993; Tomer et al., 2002). Neuroimaging studies have since brought converging results. Prefrontal, parietal and subcortical regions are most specifically involved in reactive flexibility (Fink et al., 1997; Konishi et al., 1998). Spontaneous flexibility, on the other hand, is thought to be underpinned more specifically by regions of the anterior portion of inferior frontal gyrus and the left dorsomedial thalamic nucleus (Paulesu et al., 1997).

More recently, this hypothesis gained more traction. Experimental studies have found evidence in favor of a bi-factorial structure of cognitive flexibility in adults (Tomer et al., 2002) as well as in children (Arán Filippetti & Krumm, 2020; Arán Filippetti & Richaud, 2017). Recently, a structural model of children’s cognitive flexibility, accounting for two second-order factors of cognitive flexibility representing reactive flexibility and spontaneous flexibility, has been proposed by Arán Filippetti and Krumm (2020). Reactive and spontaneous components of cognitive flexibility have also been differentially linked to several aspects of learning such as creativity, verbal skills, spatial representation, and arithmetic performance (Arán Filippetti & Krumm, 2020; Arán Filippetti & Richaud, 2017; Ebersbach & Hagedorn, 2011). Notably, Arán Filippetti and Krumm (2020) found that spontaneous, but not reactive measures of cognitive flexibility, predicted reading comprehension and writing skills in 8- to 12-year-old children. These results illustrate the relevance of separately studying the two components of cognitive flexibility, especially given their differential relationships with academic-related learning.

Furthermore, reactive flexibility and spontaneous flexibility have been shown to entail different executive demands among children. For example, in Arán Filippetti and Krumm’s study (2020), working memory only contributed to spontaneous flexibility and not to reactive flexibility. It has been assumed that reactive flexibility requires a lesser working memory load than spontaneous flexibility, since reactive shifts result from changing external contingencies whereas spontaneous changes emerge from internal inputs (Arán Filippetti & Krumm, 2020).

In sum, while most studies investigating cognitive flexibility measure it using *rule-shifting* and *switching* paradigms, evidence for a dissociation between reactive flexibility and spontaneous flexibility has indicated that these tasks may only assess the reactive component of cognitive flexibility. This highlights the need to further investigate the factorial structure of this construct using tasks allowing for the testing of reactive flexibility on the one hand (e.g., *rule-shifting* tasks), and spontaneous flexibility on the other hand (e.g., *divergent thinking* tasks) – so as to avoid the implicit assumption of a unidimensional, unique factor of cognitive flexibility underlying individuals’ performance across contexts.



## A Task-Dependent Skill?

A third line of research suggests that cognitive flexibility is far more heterogeneous than it may seem. New cognitive flexibility tasks have been designed with the aim to compare individuals' performance between different contexts, thus bringing evidence in favor of the task-dependent nature of this construct. For this purpose, participants were provided with several cognitive flexibility tasks that only differed regarding a targeted parameter. For instance, Chevalier and Blaye (2009) examined how variations of cues' transparency affected individuals' cognitive flexibility in translating external demands to the corresponding task goal. In a modified version of the DCCS, children and adults sorted cards following arbitrary or transparent changing cues. For instance, the Color rule was cued by a gray shape in the arbitrary condition, and by a multicolored shape in the transparent condition. Results showed that, while all participants could understand which sorting rule the different types of cues referred to, performances were higher with transparent cues compared with arbitrary cues. Specifically, this study showed that 5- and 6-year-old children's ability to shift between rules was significantly lower when the shifting cues imposed most cognitive demands. Similarly, a study of Blaye and Maintenant (2008) has investigated the influence of task-related conceptual knowledge on *rule-shifting* performance. They found that children succeeded in cognitive flexibility tasks involving objects categorized based on *perceptual* criteria (e.g. Color, Shape) before being able to do the same for *semantic* criteria (e.g. Taxonomic, Thematic relations), thus differentiating cognitive flexibility sorting objects based on semantic versus on perceptual relations. The authors interpreted these results as evidence for the later mastery of semantic relations between objects, compared to perceptual relations. In the same vein, using tasks which only differed in the specific stimuli to be categorized and their corresponding shifting-cues, studies have indicated that cognitive flexibility may be impacted by subtle content-related differences in the tasks, such as the kind of stimuli (e.g., either objects or creatures), or the perceptual salience of their properties (Deák & Narasimham, 2014; Deák & Wiseheart, 2015). In the same vein, other works have identified additional task-related factors impacting cognitive flexibility, including the presence of conflicting information (Blakey, Visser, & Carroll, 2016), the task-specific goal (Deák & Wiseheart, 2015; Legare et al., 2018), the complexity of shifting-cues (Chevalier & Blaye, 2009; Deák & Narasimham, 2014), and the type of shift to execute (Ravizza & Carter, 2008).

In a similar perspective, research has indicated that specific executive processes are differentially linked with cognitive flexibility depending on the task (Arán Filippetti & Krumm, 2020; Arán Filippetti & Richaud, 2017; Deák & Wiseheart, 2015). For example, in a study conducted by Deák and Wiseheart (2015), inhibitory control contributed to cognitive flexibility inferring the meaning of new words (in a *predicate-shifting* task), but not to cognitive flexibility sorting objects (in a *rule-shifting* task). What is more, it has been found that inhibitory control contributed to cognitive flexibility inferring the meaning of new words referring to objects' properties, but not to characters' properties (Deák & Wiseheart, 2015).

Besides, Deák's work (Deák, 2000, 2003; Deák & Narasimham, 2014; Deák & Wiseheart, 2015) has raised the question of whether one's ability to shift from one task to another differs depending on the activity under study. In fact, many conclusions drawn about cognitive flexibility are generalizations from *rule-shifting*. Hence, new tasks have been developed to

study cognitive flexibility inferring the meaning of words: the Flexible Induction of Meaning (FIM, Deák & Narasimham, 2014; Deák & Wiseheart, 2015) tasks. In these *predicate-shifting* paradigms, word learning is operationalized as the ability to infer the meaning of different pseudo-words based on verbal predicates preceding them. For instance: “this character *is a* [pseudo-word for character’s Species]”, “this character *lives in a* [pseudo-word for character’s Environment]”. Like *rule-shifting* tasks (e.g., the DCCS), *predicate-shifting* tasks assess cognitive flexibility in categorizing stimuli based on their relations: after having paired stimuli according to a first relevant property (e.g., same Specie), children are given new instructions asking them to pair again stimuli according to a different property (e.g., same Environment) (see Fig. 4). However, while *rule-shifting* tasks require to shift between rules (to sort stimuli), *predicate-shifting* tasks require to shift between verbal predicates (to infer the meaning of new pseudo-words). To test whether cognitive flexibility varies depending on the task, Deák and Wiseheart (2015) compared the performance of 3- to 5-year-old children in two *predicate-shifting* tasks: Flexible Induction of Meaning [Words for Objects] (FIM-Ob, Deák 2000) and the Flexible Induction of Meaning [Words for Animates] (FIM-An, Deák & Narasimham, 2014), as well as in a *rule-shifting* task: the Three Dimensional Card Sorting test (3DCCS, Deák & Wiseheart, 2015). The 3DCCS – a modified version of the DCCS – and the FIM tasks are based on the same structure. Overall, the authors reported strong correlations between children's performance on the two FIM tasks, but no links between either of FIM-Ob and FIM-An with the 3DCCS. In the same vein, Legare et al. (2018) aimed to determine whether children's cognitive flexibility is a general competence, or if it varies depending on the task domain. To address this question, the performance of 3- to 5-year-old American and South African children were analyzed in the *rule-shifting* 3DCCS and the *predicate-shifting* FIM-An tasks described above. The results showed that South African children performed lower on the *rule-shifting* task than on the *predicate-shifting* task at 5 years old, while they exhibited similar performance on these tasks at 3 years old. On the other hand, American children showed similar age-related increases in performance on both tasks between 3- to 5-year-old. The authors concluded that cognitive flexibility is not a general ability, but relies on different task-related processes and pieces of knowledge that develop differently with age. Taken together, these studies have indicated that *rule-shifting* and *predicate-shifting* are two separate skills among preschoolers (Deák and Wiseheart, 2015; Legare et al, 2018).

Thus, comparative analyses of individuals’ performance between tasks have enabled the study of cognitive flexibility across contexts. These works have shown that shifting between tasks, information or responses is more or less carried out by the individuals depending on the task to be performed, thus providing evidence for the task-dependent nature of cognitive flexibility. Interestingly, this third line of research could also account for variations of cognitive flexibility performances that have been observed between tasks imposing different executive functions demands, thus suggesting that cognitive flexibility may reflect several moderately task-dependent processes (e.g., Blakey, Visser, & Carroll, 2016; Diamond, 2013; Kim et al., 2011; Ravizza & Carter, 2008). Taken together, these studies suggest that the variability of performance from one cognitive flexibility task to another may be due to the fact that these different tasks rely on different abilities to begin with.



## The Present Study

The abundance of research testifies to the interest for the study of cognitive flexibility. However, as we have seen, different conceptions emerge in the current literature about the nature of this construct (see Fig. 1 for a schematic overview of these three main lines of research). Cognitive flexibility is either described as: (H1) a *general executive process* within cognitive control – mostly corresponding to *set shifting*, (H2) a *dimensional construct* composed of reactive and spontaneous flexibilities, or (H3) a *task-dependent skill* that varies depending on the context.

We propose that such seemingly contradictory conclusions could stem from the research frameworks under which these studies are carried out, often allowing for the testing of one postulate but not the others. Indeed, to our knowledge, no study has simultaneously compared the predictions of these three bodies of research on a single dataset. On the one hand, testing whether cognitive flexibility is a *task-dependent skill*, Deák's work (Deák & Narasimham, 2014; Deák & Wiseheart, 2015; Legare et al., 2018) has provided convincing evidence that preschoolers' cognitive flexibility performance is influenced by the context (e.g., cognitive flexibility in sorting objects differs from cognitive flexibility in inferring the meaning of words). To determine whether or not such findings replicate in older children would provide crucial insights into the robustness of these results across development. Additionally, despite recent work investigating the dimensional structure of cognitive flexibility (Arán Filippetti & Krumm, 2020), no study has, to the extent of our knowledge, simultaneously compared the predictions from the aforementioned three lines of research regarding the nature of cognitive flexibility.

In this paper, we intend to shed light on this multifaceted issue by studying the performance of first-graders in four cognitive flexibility tasks designed to tap into distinct aspects of cognitive flexibility: *rule-shifting*, *predicate-shifting*, *role-shifting*, and *divergent thinking*. The two former were meant to target reactive flexibility, while the two latter focused on spontaneous flexibility. Additionally, based on the observation that working memory and inhibition correlate with cognitive flexibility (Miyake et al., 2000) as well as on recent evidence for the contribution of processing speed to children's cognitive flexibility performance (Deák & Wiseheart, 2015), we also included three executive functions tasks (*cognitive inhibition*, *verbal working memory* and *visuomotor processing speed*) to investigate their link with cognitive flexibility.

To compare the predictions of the three competing hypotheses (summarized in Fig. 1), we propose to use correlational analysis as well as Confirmatory Factorial Analyses (CFA) on children's performance in the experiment. First, we intend to contrast those predictions by looking at the correlations between children's scores in the seven tasks. According to the *general executive process* hypothesis (H1), we should observe significant, positive correlations among the four cognitive flexibility tasks, as well as between the cognitive flexibility tasks and the executive functions tasks (Diamond, 2013). Conversely, the *dimensional construct* hypothesis (H2) predicts stronger intra-factor correlations (i.e., correlations between tasks targeting the same subcomponent of cognitive flexibility: reactive or spontaneous), than inter-factor correlations (i.e., between tasks targeting different subcomponents), in accordance with the hypothesized bi-factorial structure of this construct (Arán Filippetti & Krumm, 2020). In

addition, we expect that executive function tasks should each correlate to a similar level to tasks targeting the same subcomponent of cognitive flexibility. Finally, according to the *task-dependent skill* hypothesis (H3), we should observe neither consistent nor systematic correlations between all four cognitive flexibility tasks on the one hand, neither between the flexibility tasks and the executive functions tasks on the other hand (Deak & Wiseheart, 2015). Second, following Arán Filippetti and Krumm’s (2020) procedure, we will conduct Confirmatory Factor Analysis (CFA) to compare the fit of a “one-factor model” accounting for a unique factor of cognitive flexibility to that of a “two-factor model” assuming two correlated factors referring to reactive flexibility and spontaneous flexibility.

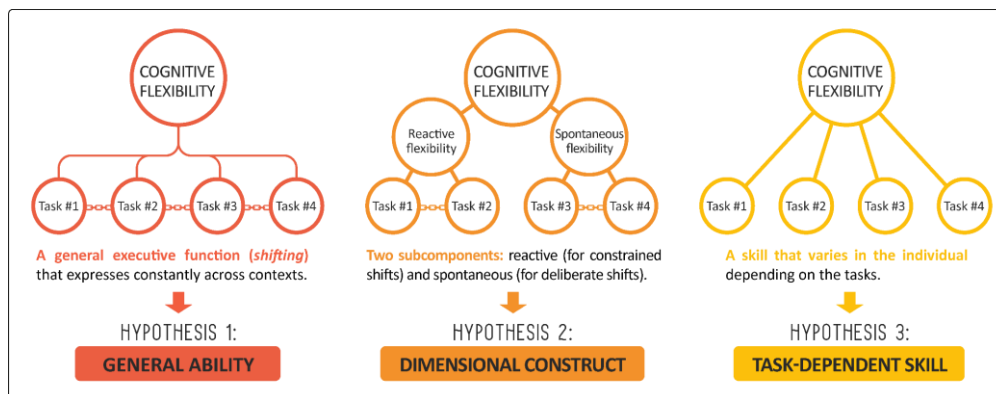


Fig. 1: Schematic representation of the three lines of research that are found in the current literature about the nature of cognitive flexibility

## Materials and Method

### Participants

Based on Deák and Wiseheart’s (2015) study among 85 children aged 3- to 5-year olds, we estimated a minimum sample size of 85 participants. We set to recruit children between 6- to 7-year olds and enrolled in first grade. We enrolled 93 first-graders in three primary public schools. All schools were located in the Parisian region. Children completed tasks individually in a quiet room in their school. After collection, data from seven children were excluded from the analyses: two children could not attend the three sessions for testing, one child failed to meet all the tasks’ inclusion criteria (i.e., the *rule-shifting* Accuracy score was out of range, see details in Materials section), three children were diagnosed with either a cognitive, sensorial, or attentional disorder, and one child was not a native French speaker. The resulting sample consisted of 86 children (mean age = 80.08 months; SD = 3.78; 40% boys).

### General Procedure

Children participated individually in a quiet room in their respective schools. They completed the seven tasks across three sessions, each lasting approximately 40 minutes, within a mean of a 10-week period. The tasks were presented in the same sequence for all participants. This was done to avoid variations in performance related to differences in task order. The sessions took

place in the afternoons so as to not interfere with scheduled morning classes. During the first of the three sessions, children were presented with the *predicate-shifting* and the *cognitive inhibition* tasks. The second session was dedicated to the two spontaneous flexibility tasks: *divergent thinking* and *role-shifting*. The last session focused on *verbal working memory*, *visuomotor processing speed*, and *rule-shifting* tasks. All sessions were video-recorded for coding verification. Parental consent for data collection and recordings was collected for each child prior to the experiment. At the beginning of each session, the experimenter reminded children of the purpose of the recording device and ensured that they were comfortable with it before starting the tasks.

Children were warned in advance that the experimenter would not give them any feedback nor help during task completion. Eye contact was avoided during response collection – whenever the task implied pointing or verbally answering – to reduce potential biases due to non-verbal suggestions from the experimenter. It was established at every other time to encourage participants. At the end of each task, children were congratulated for their participation, regardless of their actual performance. Short breaks (chatting sessions about 3 to 5 minutes) were included before starting any task.

## Materials

### Reactive Flexibility Tasks.

The two reactive flexibility tasks shared common features. Both were shifting tasks (*rule-based* and *predicate-based*) assessing participants' ability to adapt responses to changing demands (a change of sorting rule and a change of verbal predicate, respectively). Both required the effective processing of four successive stimuli's properties (one per block). The two tasks included four blocks of six trials each and provided as many response options, trials, shifts, stimuli, and their properties. They also involved the same response mode (by pointing stimuli) and were presented on a similar material (printed colorful stimuli).<sup>1</sup>

***Rule-shifting:*** We created the Four Dimension Changes Card Sorting (4DCCS) task to assess *rule-shifting* in categorizing objects. The 4DCCS is a new, modified version of the Three Dimension Changes Card Sorting (3DCCS; Deák & Wiseheart, 2015) task intended for first-graders (while the 3DCCS was created for preschoolers)<sup>2</sup>. The difference between our task and the 3DCCS lies in its difficulty. To be in line with first-graders' level, we complexified the task by adding one block (i.e., one rule change). Consequently, the 4DCCS implies more shifting and responses options compared to the 3DCCS: children had to deal with three rule shifts instead of two, and to process stimuli varying on four different properties instead of three.

---

<sup>1</sup> Since the two reactive flexibility tasks were based on the same structure – and procedure, a concern shared with Deák and Wiseheart (2015) was whether contextual similarities may raise individual response strategies across the two tasks. To minimize contextual priming due to such similarities shared between reactive flexibility tasks, the *predicate-shifting* task was the first task administered during the first session, and *rule-shifting* was assessed at the end of the third session.

<sup>2</sup> The 4DCCS was pre-tested in an independent sample to ensure that first-graders could understand the rules and complete the task (N = 14; 43% boys).

On an A3 sheet, children were presented with stimuli depicting animals with four different properties: Color, Character, Number and Orientation (see Fig. 2). Thus, for example, animals' Color was either red, green, blue, or yellow. Five target-stimuli (or *referent*), including one foil, figured on the top half of the sheet. Below the referents were six stimuli-tests (see Fig. 3). The six stimuli-test were first masked. For each trial, the experimenter revealed one of them. Thus, children were always presented with the four referent together with one stimulus to sort for each trial. Children were told to pair the stimuli-test with the referent sharing the same property depending on the current sorting rule (see Fig. 2 for an example of correct sorting of a stimulus-test for each block). First, children had to identify the referent sharing the same animals' Color with the stimuli-tests (block 1), then the same Character (block 2), third the same Number of animals (block 3), and the same animal's Orientation (block 4). Thus, for example, the stimulus-test depicting *Four Red Cows Rotated 90°* had to be paired with the leftmost referent (*One Red Pig Rotated 0°*) in block 1, and with the rightmost referent (*Two Yellow Dogs Rotated 90°*) in block 4 (see Fig.1). Blocks and trials were displayed in the same order for all children. A given property's value (e.g., *rotated 90°*) never figured on more than two stimuli-test, and any combination between two properties' values (for instance, *blue + pig*) figured only once among the six stimuli-test. Thus, trials were unambiguous as they implied only one possible correct answer. Stimuli and schematic procedure of the 4DCCS are illustrated in Fig. 3.

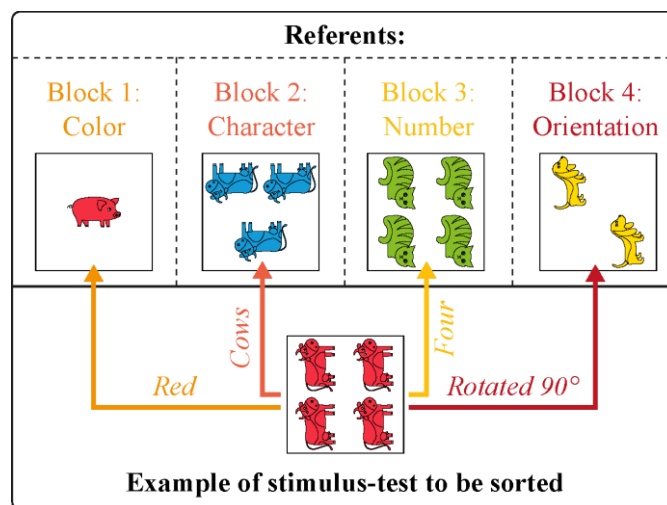


Fig. 2: The Four Dimension Changes Card Sorting (4DCCS) task – example of stimulus-test to be sorted across the four blocks. Children have to pair, one at a time, six test-cards with the referent sharing the same property according to the current sorting rule: Color (block 1); Character (block 2); Number (block 3); Orientation (block 4). A fifth referent was a foil, which is not depicted in this figure.

Based on the assumption according to which working memory activation and maintenance of a sorting rule support cognitive flexibility (Diamond, 2013), we minimized these demands. First, the need to activate a new rule was emphasized on the first trial for each block, in explicitly announcing a rule change while explaining new instructions (for example, in block 1: “let’s play the Color game, these are rules...”, see instructions into black squares in Fig. 3). Second, maintenance of sorting rules was facilitated on each subsequent trial through an implicit reminder of the current task (e.g., “in the Orientation game, can you point to the card

this one goes with?”). A demonstration phase ensured that children could recognize the stimuli’s properties (e.g., “Which color is the animal on that card? (...) and on this one?” [Experimenter pointed to the referents one by one]). All participants identified the properties as the differences between them.

We assessed *rule-shifting* by computing the ratio between *Correct Switches* and *Opportunities to Switch* (CORSWOPS score, Deák & Narasimham, 2014)<sup>3</sup>. Accuracy score, corresponding to the number of correct responses in the first block (i.e., responses on trials that did not follow a rule change) was calculated as an indicator of children’s ability to understand the task-goal. Accuracy scores under .17 (i.e., less than one out of the six correct trials on block 1) were thought to demonstrate that children were not able to follow a sorting rule categorizing objects accordingly. One child failed to reach this criterion and was excluded from the analyses. Both CORSWOPS and Accuracy scores ranged from 0 to 1.

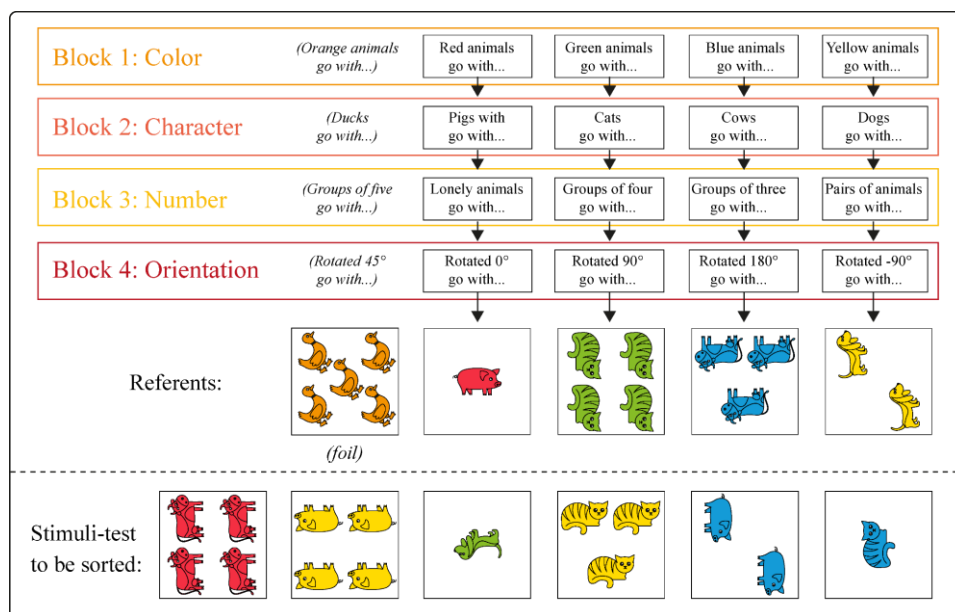


Fig. 3: 4DCCS – schematic representation of the procedure. For each block, the six stimuli-test have to be sorted one by one with the referent sharing the same property according to the current rule: block 1: Color, block 2: Character; block 3: Number; block 4: Orientation. The fifth referent is a foil, which does not share any property with the stimuli-test. The complete stimuli created for this task are available online at <https://osf.io/74gp9/>

**Predicate-shifting:** In the same vein, we created the Four Flexible Induction of Meaning for words-Animates (4FIM-An) task, to assess verbal-cue shifting inferring the meaning of words. The 4FIM-An was adapted from the FIM-An task developed by Deák and Narasimham (2014). As in the 4DCCS, we increased the difficulty by adding one block (i.e., one verbal predicate change)<sup>4</sup>. Compared to the FIM-An, the 4FIM-An task provides three predicate shifts instead of two, as well as more stimuli, with four different properties instead of three.

<sup>3</sup> CORSWOPS stands for [*Correct Switches*] : [*Opportunities to Switch*]. CORSWOPS score takes into account every *post-switch trial* (i.e., trials from blocks 2, 3, and 4) in which children give a correct response that differs from the corresponding response on the same item in the previous block. Thus, CORSWOPS does not take into account appropriate, but perseverative responses in cognitive flexibility scores (Deák & Narasimham, 2014).

<sup>4</sup> The 4FIM-An was pre-tested in an independent sample to ensure that first-graders could understand the rules and complete the task (N = 15; 33% boys).



We printed six A4 sheets with six stimuli on each. Stimuli depicted chimerical creatures standing in a strange environment, holding something unfamiliar, and looking at a bizarre object. On each sheet, a target-stimulus (or *referent*) figured at the top, with five stimuli-test (including one foil) below the referent (see Fig. 4). Each stimulus-test – except the foil – shared a unique property with the referent: either the same Character, the same Object being held by the character, the same Environment in which the character appeared to be living, or the same object it Gazes at. For each trial, children had to understand the meaning of a new PSEUDO-WORD (capitalized in this paragraph) referring to one of the referent’s properties: block 1: Character; block 2: Object; block 3: Environment; block 4: Gaze. Whenever children were presented a sheet of stimuli, they heard sentences with a new pseudo-word, as well as a verbal predicate (italicized in this paragraph) from which they had to infer which of the four referent’s properties the pseudo-word referred to. The four predicates (one per block) were: “*is a [...]*” for Character trials, “*holds a [...]*” for Object trials, “*lives in a [...]*” for Environment trials, and “*looks a [...]*” for Gaze trials. For example, when presented the sixth sheet during the first block (see Fig. 5), children were first told about the referent “this one *is a* ROLAT”. They were then asked to find the stimulus-test sharing the property the pseudo-word referred to (e.g., “Look at these five. Can you show me which one *is a* ROLAT like this one?”). Hence, children had to understand that the pseudo-word ROLAT referred to the Character depicted on the referent, and point to the stimulus-test sharing this property (in the example presented on Fig. 5, this is the leftmost stimulus on the sixth sheet in block 1). Children heard a new pseudo-word for each trial. Thus, each block involves inferring the meaning of six different pseudo-words (one per sheet) referring to the same stimuli’s property (e.g., six *Characters* in block 1; six *Objects* in block 2), cued by the same verbal predicate. The six sheets were presented one by one within a block, in the same order across the four blocks. Thus, for example, when children were presented with the sixth sheet in the second block, they heard a different verbal predicate as a new pseudo-word (“this one *holds a* PADI” – see Fig. 5). This time, they had to infer the meaning of PADI by identifying the stimulus-test depicting the creature holding the same object as the referent, thus providing a different response than they did on the first block. Meaning of pseudo-words could thus be inferred one trial after another by using the verbal predicates provided by the experimenter.

We minimized the working memory demands of the task by warning children about a change of the referents’ property to be inferred between two blocks (e.g., “Now we will learn new things we did not learn yet about the strange creatures”), and by reminding which of the referents’ properties children had to focus for each trial (e.g., after having been told about the referent “this one *holds a* PADI”, children were immediately asked “Can you show me which one *holds a* RIPOC like this one?”).

As in the 4DCCS, *predicate-shifting* corresponds to CORSWOPS score (see footnote 3 above). Accuracy scores (reflecting children’s ability to infer the meaning of words based on a verbal predicate) were computed from the proportion of correct trials in the first block. Accuracy scores under .17 were thought to demonstrate that children were not able to pursue the task’s goal. All Accuracy scores were above this ratio. Both CORSWOPS and Accuracy scores ranged from 0 to 1.

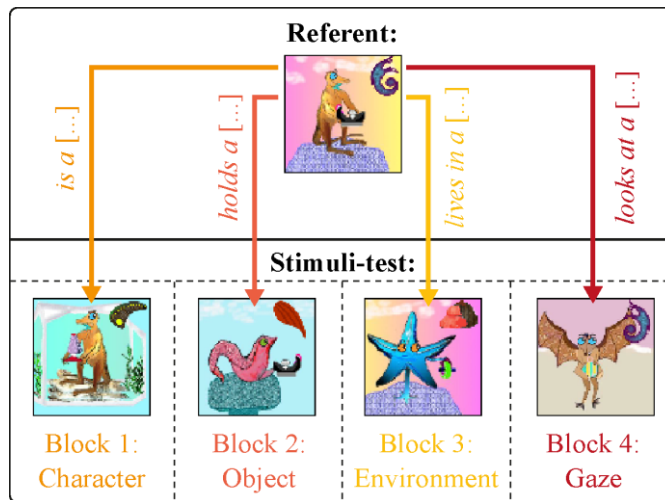


Fig. 4: The Four Flexible Induction of Meaning-Animate (4FIM-An) task – example of item. Children have to infer the meaning of different pseudo-words corresponding to the referent’s property by identifying the stimulus-test sharing the same property, based on the verbal predicate preceding the pseudo-words: Block 1: “*is a* [Character]”; Block 2: “*holds a* [Object being held]”; Block 3 “*lives in a* [Environment]”; Block 4 “*looks at a* [object Gazed at]”. A fifth stimulus-test was a foil, which is not depicted in this figure.



Fig. 5: 4FIM-An (simplified) schematic representation of the procedure. For each block, children have to infer the meaning of six new pseudo-words cued by verbal predicates: Block 1: “*is a* [Character]”; Block 2: “*holds a* [Object being held]”; Block 3 “*lives in a* [Environment]”; Block 4 “*looks at a* [object Gazed at]”. Each sheet comprises one referent and five stimuli-test (including a foil which does not share any of the referent’s properties). Unlike this simplified representation, the location of the five stimuli-test differs across the six sheets. The complete stimuli created for this task are available online at <https://osf.io/74gp9/>

### Spontaneous Flexibility Tasks.

We used two measures of spontaneous flexibility that required children to deliberately shift from one perspective or idea to another: either to produce a variety of novel ideas or to consider a role-change about the same object.

**Divergent thinking:** In the Alternative Uses Test (AUT, Guilford, 1967), children were given two minutes to come up with as many new and unusual uses for a box as they could. Following the example of Eslinger and Grattan (1993) as well as Rastelli, Greco, and Finocchiaro (2020), we used the number of different uses proposed by the children to measure spontaneous flexibility. *Divergent thinking* score corresponds to the total number of distinct categories that were identified in the children's answers (ranging from 0 to 6 in our sample).

Response rating was performed by the first author, in accordance with the quotation procedure recommended by Reiter-Palmon, Forthmann, and Barbot (2019). For example, proposing to use a box as a container (e.g., to store objects) and then as a support (e.g., to draw or write on) was rated as using two distinct categories.

**Role-shifting:** Given the debate regarding whether *verbal fluency* tasks should be used as a measure of reactive flexibility or spontaneous flexibility (e.g., Diamond, 2013; Eslinger & Grattan, 1993; Kavé, Kigel, & Kochva, 2008; Munakata, Snyder, & Chatham, 2012), we elected not to rely on these tasks. Instead, we opted to design a new spontaneous flexibility task fitting the needs of our study. The Spontaneous Role-Shifting task (SRS) was created to target spontaneous flexibility through *role-shifting*<sup>5</sup>. We define a “spontaneous role shift” as a deliberate change from an initial representation – or point of view – about an object, to a different one mobilizing conceptual knowledge (for instance, spontaneously consider that a *dad* is his son’s *father*, but also his own fathers’ *son*).

Children were presented twelve printed sheets (six target-items and six distractors). Each sheet was composed of three drawings (most often depicting a character, an animal, or an object), as well as a fill-in-the-blanks sentence stating a relationship between two entities (e.g., “... *eats* ...”; see Fig. 6). Children were instructed to verbally state a *realistic* sentence by filling the blanks with two of the labels written below the three drawings, as many times as possible. For target-items, the sentence could be completed twice (e.g., “the wolf *eats* the sheep” and “the sheep *eats* the grass”). Three distractor-items provided a sentence that should only be realistically completed once, and three others a sentence that could not be completed at all with the drawing’s labels (see red crosses in Fig. 6). Children were asked to complete the sentences as many times as possible, without being told how many times each sentence could be filled.

The test was preceded with a demonstration involving three sample sheets, one for each type of item (*no-sentence* distractor, *1-sentence* distractor, *2-sentence* target – demonstration items are displayed in Fig. 6). This phase also enabled the experimenter to estimate the reading level of children. At the beginning of each trial, the three drawings were first masked. The experimenter read the fill-in-the-blanks sentence, scanning it with her finger. Then, the three

---

<sup>5</sup> The SRS was pre-tested in an independent sample to ensure that first-graders could understand the rules and complete the task (N = 10; 3 boys).

drawings were revealed. Their labels (e.g., “the wolf”, “the sheep”, and “the grass”; see Fig. 5) were orally stated before asking children whether or not the sentence could be completed. On target items, using the same sentence structure twice requires to spontaneously shift from a first point of view to another one, since one of the labels had to be reused in a different position in the sentence (i.e., in the example presented in Fig. 6: “the sheep” could have two different roles). To ensure that children did not simply propose every possible combination of labels (thus producing random and unrealistic relationships between two stimuli), we checked that the total number of unjustified incorrect matches was lower than the total number of correct matches. All children meet this inclusion criterion. *Role-shifting* score corresponds to the total number of correct shifts on target items (ranging from 0 to 6).

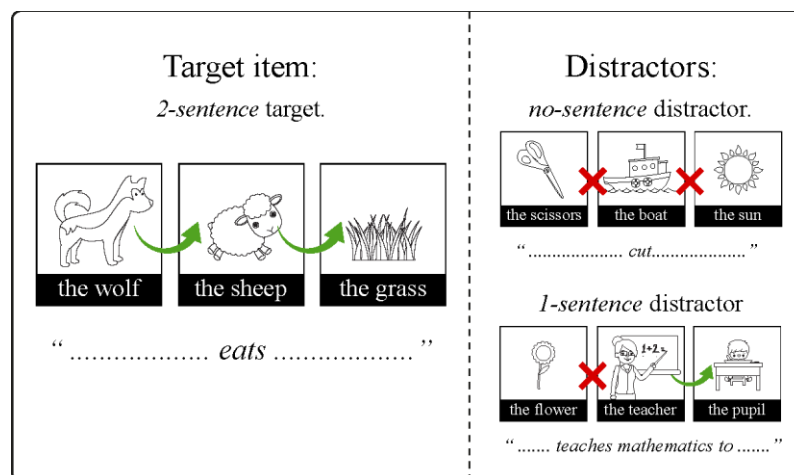


Fig. 6: The Spontaneous Role-Shifting task (SRS) task – demonstration items. Whenever possible, children have to fill-in sentences with two drawing labels. For target items, the sentence should be completed twice. Unlike these simplified examples, the location of the drawings and their relationship are not spatially congruent. The complete stimuli created for this task are available online at <https://osf.io/74gp9/>

### Executive Functioning Tasks.

**Verbal working memory:** We used the BREV task (Billard et al., 2000) as a measure of children’s verbal working memory. Participants were instructed to recall, in reverse order, sequences of two (e.g., 4-9) to six numbers (e.g., 8-5-2-9-4-6) enumerated by the experimenter (i.e., the last heard number had to be recalled first). Following Billard et al. (2000) procedure, one extra digit was added every two trials. Thus, the task began with two sequences of two digits, followed by two sequences of three digits, and so on. The task stopped whenever participants failed to correctly recall two series with the same number of digits. Children failing to recall in reverse order two digits during a training phase would be excluded from the analyses. All children correctly answered training items. Backward *verbal working memory* span score corresponds to the largest correctly retrieved number of digits sequence (ranging from 2 to 6 in our sample).

**Visuomotor processing speed:** We used the NEPSY-II processing speed task (Korkman, Kirk, & Kemp, 2012), as a measure of *visuomotor processing speed*. Children were

successively presented with two A3 sheets with numerous drawings (e.g., clothes, fruits, toys). Both sheets displayed the same drawings: either aligned on a grid on the first sheet, and misaligned in the second one. For each sheet, children were asked to mark as many clothes as possible in 45 seconds. This task involves the fast processing of perceptive and semantic information to correctly identify whether or not an object belongs to a relevant category (i.e., clothes). A training phase ensured that children understood and could complete the task. *Visuomotor processing speed* score corresponds to the total number of marked clothes on the two trials, minus the errors made by marking irrelevant drawings. Negative scores are replaced by '0'. The resulting scores may range from 0 to 96.

***Cognitive inhibition:*** We relied on the NEPSY-II subtest (Korkman, Kirk, & Kemp, 2012) to assess inhibitory control. Children were shown an A4 sheet depicting 40 shapes (circles or squares), followed by another sheet depicting 40 arrows (pointing up or down). The two sheets were processed sequentially in two stages. In the first stage, children were asked to pronounce out loud the name of the shapes on the first sheet (i.e., “circle” or “square”), and the direction of the arrows on the second sheet (i.e., “up” or “down”), within a 180-second time period for each. In the second stage, they were instructed to reverse the names of the shapes on the first sheet (i.e., saying “circle” when they saw a square – and vice versa), and the direction of the arrows on the second sheet (i.e., saying “up” when they saw an arrow pointing down – and vice versa), within a 240-second time period for each. *Cognitive inhibition* was assessed during the second stage, which involves verbal inhibition of automatic associations. *Cognitive inhibition* score corresponds to the inverted value of the sum of uncorrected and self-corrected errors made on the two sheets (ranging from -39 to -1 in our sample).

## Results

The complete analyses procedure can be consulted at <https://osf.io/74gp9/>

### Correlational Analyses

After scoring participants' responses, we looked at the correlations between the seven measures in order to contrast our three hypotheses. Table 1 shows descriptive statistics and correlations among all measures.

***Correlation among cognitive flexibility tasks.*** Significant positive correlations emerged between five out of the six pairs of cognitive flexibility tasks, with homogeneously moderated strengths ( $r$  coefficients ranged from .23 to .33). The correlation between the reactive *predicate-shifting* and the spontaneous *role-shifting* tasks did not reach significance level – despite being remarkably close ( $p = .053$ ). The two reactive flexibility tasks (*predicate-shifting* and *rule-shifting*) were positively correlated with each other ( $r = .29$ ,  $p = .007$ ,  $CI_{95\%} = [.08, .47]$ ), and so were the two spontaneous flexibility tasks (*role-shifting* and *divergent thinking*;  $r = .27$ ,  $p = .010$ ,  $CI_{95\%} = .07, .46$ ). The correlations among reactive flexibility and spontaneous flexibility task pairs showed similar effect sizes and 95% confidence intervals. Interestingly,



most of the reactive flexibility and spontaneous flexibility scores were also correlated to each other (with the exception of the reactive *predicate-shifting* and the spontaneous *role-shifting* tasks), with  $r$  coefficients ranging from .23 to .31,  $ps < .05$ .

Table 1: Correlations among measures ( $r$  coefficient with 95% confidence intervals).

	Reactive flexibility		Spontaneous flexibility		Executive functions		
	Pred.-shift 1	Rule-shift 2	Role-shift 3	Div. think 4	Inhib. (inv) 5	vmSpeed 6	vWM 7
1 Pred.-shift							
2 Rule-shift	<b>.29***</b> [.08, .47]						
3 Role-shift	.21 [-.00, .40]	<b>.31**</b> [.11, .49]					
4 Div. think	<b>.33**</b> [.13, .51]	<b>.23*</b> [.02, .42]	<b>.27*</b> [.07, .46]				
5 Inhib. (inv)	.11 [-.10, .32]	.05 [-.16, .26]	.00 [-.21, .21]	.02 [-.20, .22]			
6 vmSpeed	.21 [-.00, .40]	<b>.42***</b> [.23, .58]	<b>.45***</b> [.26, .60]	<b>.27*</b> [.07, .46]	-.10 [-.30, .12]		
7 vWM	.05 [-.16, .26]	<b>.22*</b> [.01, .41]	.15 [-.06, .35]	-.06 [-.27, .15]	.15 [-.06, .35]	.04 [-.17, .25]	
Mean	0.74	0.76	3.42	2.13	-9.06	45.13	3.20
SD	0.23	0.20	1.76	1.42	7.36	7.64	0.94
Skewness	-0.73	-1.51	-0.30	0.59	-1.46	0.07	0.61

Pred.-shift = *predicate-shifting*; Rule-shift = *rule-shifting*; Role-shift = *role-shifting*; Div. think = *divergent thinking*; vmSpeed = *visuomotor processing speed*; vWM = *verbal working memory*; Inhib. (inv) = *cognitive inhibition (inverted score)*  
 $*p < .05$ ;  $**p < .01$ ;  $***p < .001$

**Correlations between cognitive flexibility and executive functions tasks.** Overall, neither *cognitive inhibition*, *verbal working memory*, nor *visuomotor processing speed* showed a consistent correlational pattern with any measure of cognitive flexibility. First, *cognitive inhibition* did not correlate with any of the four cognitive flexibility tasks. In contrast, *verbal working memory* scores were found to moderately correlate with *rule-shifting* ( $r = .22$ ,  $p = .040$ ,  $CI_{95\%} = [.23, .58]$ ), but not with *predicate-shifting*, *role-shifting*, nor *divergent thinking* performance ( $ps \geq .17$ ). Finally, the *visuomotor processing speed* task was related to *rule-shifting* ( $r = .42$ ,  $p = .000$ ,  $CI_{95\%} = [.23, .58]$ ), *role-shifting* ( $r = .45$ ,  $p = .000$ ,  $CI_{95\%} = [.26, .60]$ ), and *divergent-thinking* tasks ( $r = .27$ ,  $p = .010$ ,  $CI_{95\%} = [.07, .46]$ ); however *visuomotor processing speed* and *predicate-shifting* were independent of each other.

**Discussion.** Under the *general executive process* hypothesis (H1), we expected correlations between all the seven tasks. Yet, not all the cognitive flexibility tasks were correlated with each other, and, additionally, none of the four cognitive flexibility tasks was found to systematically

correlate with the three executive functions tasks. On the other hand, the results did not make it possible to state whether cognitive flexibility is a *dimensional construct* composed of reactive flexibility and spontaneous flexibility (H2). In accordance with H2 predictions, both reactive flexibility and spontaneous flexibility tasks pairs were correlated with each other, and both were differently associated with the executive functions tasks. However, several inter-factor correlations were stronger than intra-factor correlations, thereby contradicting the assumption that most of the shared variance between two cognitive flexibility tasks would be accounted for by the underlying factor to which they belong. Furthermore, results showed specific associations between spontaneous flexibility tasks, but not reactive flexibility tasks, with executive functions. While these results may have suggested a unique factor of cognitive flexibility, this postulate was contradicted due to the lack of correlation between two of our flexibility tasks. Overall, these findings appear to provide support for the *task-dependent* nature of cognitive flexibility (H3).

### **Confirmatory Factorial Analyses (CFA)**

While correlational analyses provided valuable insights into our research question, Confirmatory Factorial Analyses should enable further investigation of a factorial structure within cognitive flexibility. As extremely low or extremely high scores may bias CFA parameters estimate, we used Tukey's method to calculate the upper and lower bounds defining the range outside of which a score might introduce bias to our analyses and should be treated as an outlier. Outlier analysis revealed four extremely low *rule-shifting* scores, one extremely low *predicate-shifting* score, as well as four extremely low scores among *cognitive inhibition* scores. These scores were removed from the dataset for CFA. The outlier-free CFA sample included N = 77 children aged 6- to 7-year-old (mean age = 80.13 months; SD = 3.74; 36% boys).

We checked for the suitability of the sample for CFA analyses using Bartlett's test of sphericity ( $\chi^2_{(6)} = 28.574$ ,  $p < .001$ ) and the Kaiser, Meyer, Olkin test (KMO = .67). The dependent variables derived from univariate normality, and multivariate normality (p-values for Mardia's multivariate skewness and kurtosis coefficients were  $p = .012$  and  $p = .466$ , respectively). Since normality assumption violations may bias parameter estimates, thus potentially affecting the true values of fit indices, we used the Maximum Likelihood Robust estimator (MLR – a robust estimator that relaxes the assumption of multivariate normality for ordinal data, Yilmaz, 2019) to estimate the parameters of the models.

Table 2: Fit indices for the “one-factor model” and the “two-factor model

Model	Model fit indices									Comparative fit indices	
	$\chi^2$ <sup>(a)</sup>	df $\chi^2$	p $\chi^2$	CFI	TLI	SRMR <sup>(a)</sup>	RMSEA	CI90% <sub>RMSEA</sub> A	P <sub>RMSEA</sub> A	BIC	AIC
One-factor	1.48	2	.48	1.00	1.00	.03	.00	[.00, .21]	.53	501.34	482.59
Two-factor	0.32	1	.57	1.00	1.00	.01	.00	[.00, .25]	.61	504.62	483.53

*Model fit indices*: CFI = Comparative Fit Index; TLI = Tucker Lewis fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual<sup>(b)</sup>; *Comparative fit indices*: BIC = Bayesian Information Criterion; AIC = Akaike Information Criterion. Low  $\chi^2$  statistic, CFI and TLI greater than .95, RMSEA value below .06, and upper values of CI90% for RMSEA below .05, are indicators of good fit. CFI and TLI greater than .98, SRMR and RMSEA values below .05; are indicators of excellent fit. Lower BIC and AIC values indicate a better fit of one model above another.

<sup>(a)</sup> As models are non-nested,  $\chi^2$  and SRMR were only interpreted as *model fit* indices

Following the statistical approach used by Arán Filippetti and Krumm’s (2020), we analyzed the fit of two models assuming either a unidimensional (“one-factor model”) or a bifactorial (“two-factor model”) structure of cognitive flexibility. The one-factor model accounted for a unique factor of cognitive flexibility underlying all cognitive flexibility scores: *rule-shifting*, *predicate-shifting*, *role-shifting*, and *divergent thinking* (see Fig. 7 for a schematic representation of the one-factor model). The two-factor model freely estimated the correlation between two factors corresponding to reactive flexibility and spontaneous flexibility. Two variables loaded on each factor: *rule-shifting* and *predicate-shifting* were used as indicators of reactive flexibility, and *role-shifting* and *divergent thinking* were indicators of spontaneous flexibility (see Fig. 8 for a schematic representation of the two-factor model). We specified the two models using the *standardized variance method*, so as to freely estimate the standardized factor loadings  $\Lambda$  (the estimated parameters for each model are indicated in Fig. 7 and Fig. 8). It should be noted that the total number of dependent variables (i.e., the four measures of cognitive flexibility) did not allow the absolute comparison of fit between two *nested models* (i.e., models which are based on the same architecture). The one-factor model was the simplest model (thus, the most parsimonious) since it explained data with a fewer number of free parameters (see bolded blue values in Fig. 7 and Fig. 8).

*Model fit* indices were  $\chi^2$  statistic, Comparative fit index (CFI), Tucker Lewis fit Index (TLI), Root Mean Square Error of Approximation (RMSEA), upper value of the 90% Confidence Interval (CI<sub>90%</sub>) for the RMSEA, and Standardized Root Mean Square Residual (SRMR). Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were used as *comparative fit* indices.

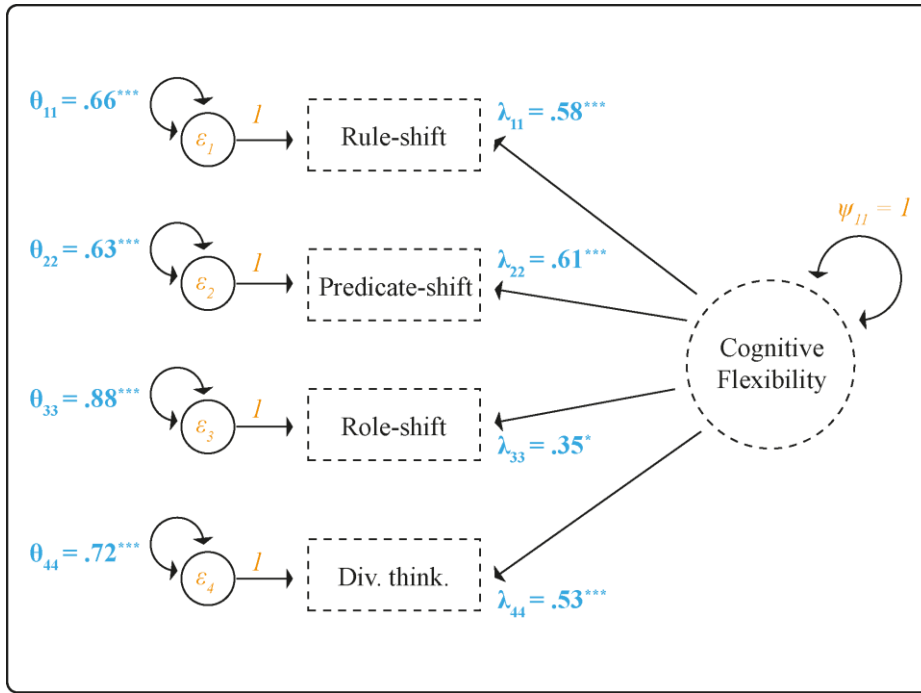


Fig. 7: Estimated one-factor model. Estimated parameters (bolded blue values) are factor loadings ( $\Lambda$ ) and residual variances ( $\Theta_{\epsilon}$ ). Fixed parameters (italicized orange values) are factor variances ( $\psi$ ) and the coefficients of residuals ( $\epsilon$ ). Dashed circle represents the latent variable; dashed squares represent dependent variables. Single-headed arrows show direction of assumed causal influence; double-headed arrows represent covariances.

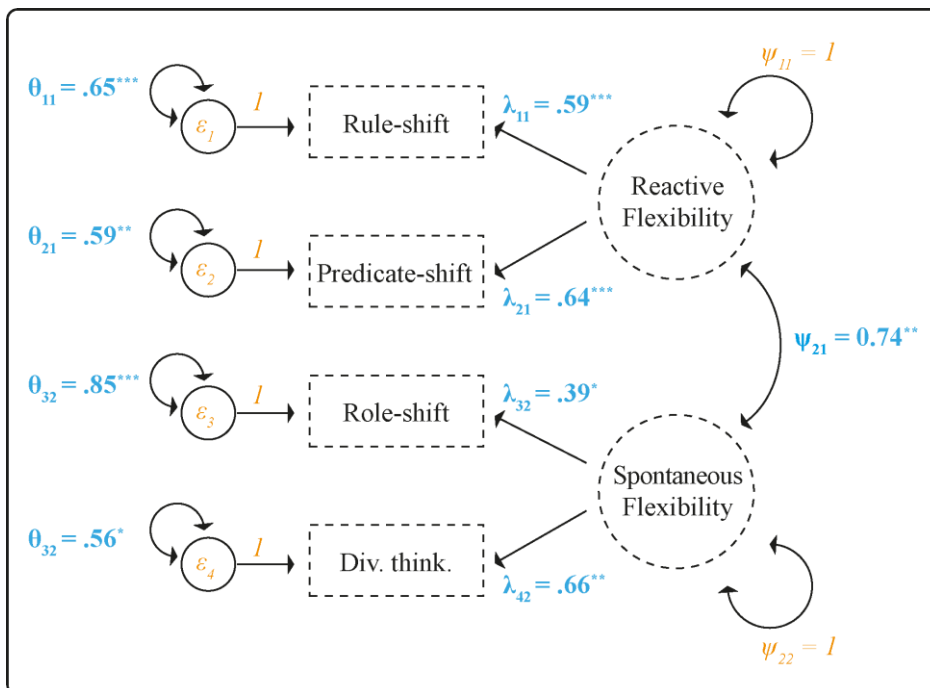


Fig. 8: Estimated two-correlated-factor model. Estimated parameters (bolded blue values) are factors covariance ( $\Psi$ ), factor loadings ( $\Lambda$ ), and residual variances ( $\Theta_{\epsilon}$ ). Fixed parameters (italicized orange values) are factor variances ( $\psi$ ) and the coefficients of residuals ( $\epsilon$ ). Dashed circles represent latent variables; dashed squares represent dependent variables. Single-headed arrows show direction of assumed causal influence; double-headed arrows represent covariances.

The CFA results are presented in Table 2. The *model fit* indices overall indicated an excellent fit for both the one-factor and the two-factor models, with CFI and TLI above .98, SRMR below .05, and non-significant RMSEA close to zero. Accordingly, low and non-significant  $\chi^2$  test values indicated a good fit to the data for the two models. It should be noted, however, that the upper values of CI<sub>90%</sub> for RMSEA were much higher than .05 for both the one-factor model (CI<sub>90%</sub>RMSEA<sub>(one-factor)</sub> = [.21]) and the two-factor model (CI<sub>90%</sub>RMSEA<sub>(two-factor)</sub> = [.25]). These results suggest that the true RMSEA values could be higher than the estimated values in our findings (both RMSEA were found to be close to zero, see Table 2). Consequently, some degree of precaution should be taken interpreting RMSEA fit indices<sup>6</sup>. In both models, the estimated standardized factor loadings ( $\Lambda$ ) revealed significant positive relation between all cognitive flexibility variables and their respective underlying factors (see  $\Lambda$  values in Fig. 7 and Fig. 8). In other words, in the one-factor model, each cognitive flexibility variable significantly loaded onto a single factor of cognitive flexibility. Meanwhile, in the two-factor model, these variables were distinctly associated with either reactive flexibility or spontaneous flexibility. Accordingly, the estimated  $\Psi$  parameter for the correlation between reactive flexibility and spontaneous flexibility on the two-factor model indicated appropriate discriminant validity between the two factors ( $\Psi_{21} = 0.74 < .85, p = .001$ ).

TLI and CFI indices did not provide insight into which model fit best, as they were equal to 1 for both models. However, *comparative fit* indices (see Table 2) were slightly lower for the one-factor model compared to the two-factor model (BIC<sub>(one-factor)</sub> = 501.34 < BIC<sub>(two-factor)</sub> = 504.62, and AIC<sub>(one-factor)</sub> = 482.59 < AIC<sub>(two-factor)</sub> = 483.53), thus indicating a slightly better fit for the one-factor model.

**Discussion.** Confirmatory Factorial Analyses indicated that, based on our data, cognitive flexibility was equally likely to be structured as either a unidimensional or a bi-factorial construct, since the overall *model fit* indices indicated that both the one-factor and the two-factor models adequately fit the data. On the other hand, while *comparative fit* indices (BIC and AIC) would have favored the two-factor model over the one-factor model, it should be noted that the one-factor model was the simplest model, and therefore the more parsimonious. Taken together, those results indicate that the current data do not definitively favor one model over the other, which leaves open the question of the dimensional nature of cognitive flexibility (H2).

## Discussion

In this study, we set out to bring empirical evidence among 6- to 7-year-old children to contrast the three main hypotheses in the literature regarding the nature of cognitive flexibility. As previously stated, cognitive flexibility has alternatively been described as a generalized executive process mostly corresponding to *set shifting* (Diamond, 2013), as a bi-factorial

---

<sup>6</sup> Post-hoc analyses have confirmed that high pRMSEA values may be due to the current sample size, providing valuable insights for future research directions (these analyses procedure can be consulted at <https://osf.io/74gp9/>). These directions may involve analyzing the one-factor and two-factor models in a larger sample. It is worth noting that CFI, TLI, and SRMR *model fit* indices used in the present CFA analyses are robust for accurately estimating the fit of models with current sample size.



construct composed of reactive flexibility and spontaneous flexibility (Eslinger & Grattan, 1993), or as a skill that varies across contexts (Deák, 2003). What can we conclude regarding these three hypotheses based on the results?

### **Is cognitive flexibility a General Executive Process?**

Within the scope of executive functioning, cognitive flexibility has been described as a generalized ability allowing for adaptation of thoughts and behaviors in any given context (Diamond, 2013; Miyake et al., 2000; Zelazo & Frye, 1998). Contrary to what would have been expected under this hypothesis, our findings indicate that children performing well in one cognitive flexibility task did not systematically perform well in others. Specifically, *predicate-shifting* and *role-shifting* scores were not related, thus suggesting that these two tasks assessed different constructs. Therefore, the significant correlations among the five others pairs of cognitive flexibility tasks cannot be interpreted as evidence for a generalized ability underlying performance across these contexts.

While according to this view, executive processes should have supported all cognitive flexibility performance (e.g., Zelazo & Frye, 1998), in our experiment neither of the executive functions tasks was found to correlate with the four cognitive flexibility tasks. Rather, our findings indicated that cognitive flexibility and executive functions were mostly independent from each other. Specifically, *cognitive inhibition* was not correlated with any cognitive flexibility task, *verbal working memory* was only found to correlate with *rule-shifting*, and *visuomotor processing speed* was related with *rule-shifting*, *role-shifting*, and *divergent thinking*, but not with *predicate-shifting*. The fact that *cognitive inhibition* was not linked with any measure of cognitive flexibility was particularly noteworthy, as *rule-shifting* tasks have often been linked to inhibitory control in the broad literature on executive functions. Under this view, *set shifting* (the shift between two tasks or stimuli) is mostly thought to rely on the ability to suppress conflicting responses or information (Diamond, 2013). Yet, our results, supported by previous findings from Deák and Wiseheart (2015) showing that cognitive flexibility in *rule-shifting* and *predicate-shifting* is not related to *cognitive inhibition* among preschoolers, suggest that different *rule-shifting* tasks involve specific executive processes depending on task characteristics.

Taken together, these results do not support the existence of a generalized ability underlying flexible behaviors across various tasks or contexts (H1). Additionally, these results emphasize why it may be crucial to study cognitive flexibility using various tasks assessing a wide range of flexible behaviors, instead of focusing on a single rule-shifting test as a measure of cognitive flexibility as a whole.

### **Is cognitive flexibility a dimensional construct?**

Our study also investigated whether our results could have accounted for a bi-factorial structure of cognitive flexibility (H2). Correlational analyses were equivocal regarding this postulate.

In line with H2 predictions, within-group correlations were observed: the reactive flexibility tasks (*rule-shifting* and *predicate-shifting*) were correlated with each other, as were the

spontaneous flexibility tasks (*role-shifting* and *divergent thinking*). Of note, these associations were not specific between two tasks targeting the same factor (intra-factor correlations), as most of the reactive and spontaneous flexibility tasks were also correlated to each other (inter-factor correlations). Since reactive flexibility and spontaneous flexibility are thought to represent two distinct but intercorrelated subcomponents of a first-order factor of cognitive flexibility (Arán Filippetti & Krumm, 2020; Eslinger & Grattan, 1993), these findings may have accounted for a two-factor structure of this construct.

However, H2 posits stronger intra-factor correlations than inter-factor correlations. In other words, it predicted higher correlations among tasks underlying the same component of cognitive flexibility than for task pairs composed of one reactive and one spontaneous task. Yet, the former were sometimes weaker than the latter. Besides, while both spontaneous flexibility tasks correlated with *visuomotor processing speed*, reactive flexibility tasks did not show specific associations with executive functions: only *rule-shifting* – but not *predicate-shifting* – correlated with *visuomotor processing speed* and *verbal working memory*. Thus, the two components showed no specific associations with executive functions, while it was assumed that two tasks belonging to the same cognitive flexibility component should involve common joint processes. These results may have suggested a unidimensional structure of cognitive flexibility. However, the non-significant correlation between the *predicate-shifting* reactive flexibility task and the *role-shifting* spontaneous flexibility task challenges both the postulate of a unique factor, and that of two correlated factors of cognitive flexibility. .

To further investigate this issue, and in line with Arán Filippetti and Krumm's procedure (2020), we compared the fit of two models assuming either a unidimensional, or a bi-factorial organization of this construct. In accordance with the correlational analysis, the CFA outputs revealed conflicting evidence favoring neither of the two accounts for the dimensional structure underlying cognitive flexibility. While both models provided adequate fits, neither clearly emerged as superior. Although *comparative fit* indices (BIC and AIC) suggested a marginally better fit for the two-factor model, the principle of parsimony leans us towards the simpler one-factor model without conclusively favoring it.

Thus, these results leave open the question of the most accurate dimensional structure of cognitive flexibility. The factorial organization of cognitive flexibility in children is a relatively recent research question which deserves further scrutiny. Our results suggest that future studies using a higher number of cognitive flexibility tasks targeting reactive and spontaneous flexibility may be able to provide conclusive evidence regarding the absolute comparison of fit between the two models. However, beyond the dimensional structure of cognitive flexibility, the predictions from the third line of research (H3) remain.

### ***Is cognitive flexibility a task-dependent skill?***

To explore the hypothesis according to which cognitive flexibility is a context-dependent skill, children were provided with four distinct cognitive flexibility tasks assessing their performance across different situations. The fact that correlations among the cognitive flexibility tasks were inconsistent suggests that variations in task characteristics led to differences in performance. Rather than uniformly strong correlations, the moderate strength of the correlations indicated

that while there might be a common underlying ability accounting for some of the shared variance, each cognitive flexibility task primarily involved unique, task-specific processes. Specifically, the lack of a significant correlation between the *predicate-shifting* and the *role-shifting* tasks directly pointed to cognitive flexibility being a task-dependent skill. This suggests that the cognitive flexibility required to infer the meaning of words following a verbal predicate change, and the cognitive flexibility necessary to shift from an initial representation to another by mobilizing conceptual knowledge, may constitute two separate skills. Along with these findings, all the executive functions tasks (*cognitive inhibition*, *verbal working memory*, and *visuomotor processing speed*) presented inconsistent links with cognitive flexibility, thus providing additional evidence that the cognitive flexibility tasks mostly elicited specific, task-dependent processes. Thus, the collected data appears to support the predictions outlined in H3 the most.

## Exploring the task-dependency

In light of the evidence for the task-dependent nature of cognitive flexibility highlighted in the analyses, further conclusions regarding the factorial structure of cognitive flexibility can be drawn. CFA analyses indicated that both a one-factor model and a two-factor model provided a good fit to the data. They did not, however, point to one of the models as a clearly better fitted model. On top of that, we might also want to question the notion that either of these models matches the true underlying structure of cognitive flexibility. An alternative hypothesis can be put forward, supported by Ionescu's proposal (2012), that cognitive flexibility is a fully task-dependent construct, which does not fall under any factorial structure. Indeed, Ionescu (2012) has proposed to define cognitive flexibility as a *property* of the cognitive system emerging from the interacting processes involved in a given flexible behavior (e.g., flexible categorization, divergent thinking, flexible word learning, flexible problem solving), rather than a static, structurally-situated cognitive *process* (e.g., *set shifting*). Under this view, the correlations between cognitive flexibility tasks in our experiment may have reflected transversal processes elicited by these tasks, the recruitment of which would have accounted for the variations in cognitive flexibility performance.

According to Ionescu (2012), research should investigate situations requiring cognitive flexibility (with the cognitive system as a whole being more or less flexible depending on each context). This underscores the importance of developing and evaluating tasks and paradigms that examine cognitive flexibility across different contexts and domains, to more fully grasp how context-specific this construct really is.

## Implications and Contributions

Promising indications for the validity of our two adapted *rule-shifting* (4DCCS) and *predicate-shifting* (4FIM-An) tasks, as well as our original *role-shifting* task (SRS) in measuring cognitive flexibility were provided by CFA analyses, showing that these three tasks significantly loaded on cognitive flexibility in both factorial models (see Results section). Additional valuable insight for the measurement validity of these tasks have been provided

with post-hoc analyses investigating the factor loadings of *predicate-shifting*, *rule-shifting*, and *role-shifting* on cognitive flexibility, relative to *divergent thinking* (these analyses procedure can be consulted at <https://osf.io/74gp9/>). As such, we believe that the new tasks developed and discussed in this study offer a significant contribution to the understanding of cognitive flexibility, as they may be used in further studies to investigate the different forms of cognitive flexibility among 6- to 7-year olds. Additionally, the analyses of the collected data also made it possible to gather new insights on the current debate as well as to highlight promising future research directions.

Some of our observations diverge from those of previous studies carried out in children, prompting further analysis to understand the origin of these differences. Arán Filippetti and Krumm's study (2020), for instance, have provided evidence for a bi-dimensional organization of cognitive flexibility among children aged 8- to 12-year-old. By contrast, our results did not provide evidence that a bi-factorial structure best accounted for cognitive flexibility of 6- to 7-year-old children, compared to a unifactorial structure. Another divergence concerns the relationship between rule-based cognitive flexibility and predicate-based cognitive flexibility. Previous studies with younger children found no correlation between the performance on *rule-shifting* and *predicate-shifting* tasks – thus suggesting that these two tasks assess distinct skills (Deák & Wiseheart, 2015; Legare et al., 2018). In contrast, our results indicate that the performance of first-graders in *rule-shifting* and *predicate-shifting* tasks are – to some extent, correlated. Finally, a third line of divergence related to the relationship between cognitive flexibility and executive functions. While Deák and Wiseheart (2015) found that *verbal working memory* was linked with *predicate-shifting* among 3- to 5-year olds, we did not find this association in our cohort. Conversely, while in our study *visuomotor processing speed* was correlated with *rule-shifting*, but not with *predicate-shifting* performance, the opposite pattern was found by Deák and Wiseheart (2015). These divergent results between studies regarding the *rule-shifting* and the *predicate-shifting* tasks are particularly noteworthy, given that we adapted the rule-based 4DCCS and the verbal predicate-based 4FIM-An cognitive flexibility tasks used in this experiment, from the previous 3DCCS and FIM-An tasks (designed by Deák & Wiseheart, 2015; and by Deák & Narasimham, 2014, respectively – see Materials section), intended for a younger population.

Several hypotheses can be put forward regarding the origin of these discrepancies. One possibility is that the age difference between the populations tested in the aforementioned studies and our sample account for the divergent results, due to the cognitive processes involved in these tasks following different developmental trajectories. Another interpretation would stem from slight differences between the specific tasks used in the experiments to assess cognitive flexibility which could, according to Ionescu (2012), account for the variation of performance. The conclusions drawn from our study, supported by previous findings (e.g., Deák & Wiseheart, 2015), lend support to this assumption, as our results clearly showcased that different cognitive flexibility tasks tap into distinct aspects of task-dependent skills.

Because task designs and paradigms currently vary widely from one study to another, interpreting and generalizing results is difficult as it stands. Consequently, and as long as we do not have a finer conceptual and empirical understanding of the nature of cognitive flexibility, we can only speculate whether the same construct is targeted across experiments. In this sense, it seems crucial that experimental research should pursue its efforts to further

study the task-dependent nature and the dimensional structure of cognitive flexibility, carefully designing or selecting tasks according to their domain and content features, while considering the cognitive processes that underlie them.

## References

- Arán Filippetti, V., & Krumm, G. (2020). A hierarchical model of cognitive flexibility in children: Extending the relationship between flexibility, creativity and academic achievement. *Child Neuropsychology*, 26(6), 770-800. <https://doi.org/10.1080/09297049.2019.1711034>
- Arán Filippetti, V., & Richaud, M. C. (2017). A structural equation modeling of executive functions, IQ and mathematical skills in primary students : Differential effects on number production, mental calculus and arithmetical problems. *Child Neuropsychology*, 23(7), 864-888.
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5 : Changes and correlates. *Developmental Review*, 29(3), 180-200. <https://doi.org/10.1016/j.dr.2009.05.002>
- Billard, C., Vol, S., Livet, M., Motte, J., Vallee, L., Gillet, P., & Marquet, T. (2000). Batterie rapide d'évaluation des fonctions cognitives (BREV). *Editions Signes*.
- Blakey, E., Visser, I., & Carroll, D. J. (2016). Different executive functions support different kinds of cognitive flexibility: Evidence from 2-, 3-, and 4-year-olds. *Child development*, 87(2), 513-526.
- Blaye, A., & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers : The development of flexibility in categorization choices. *British Journal of Developmental Psychology*, 19(3), 395-412. <https://doi.org/10.1348/026151001166173>
- Blaye, A., & Maintenant, C. (2008). Développement de la flexibilité catégorielle de 3 à 8 ans : Rôle des aspects conceptuels. *L'Année psychologique*, 108(4), 659-698.
- Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability : Inhibition, Switching, and Working Memory. *Developmental Neuropsychology*, 19(3), 273-293. [https://doi.org/10.1207/S15326942DN1903\\_3](https://doi.org/10.1207/S15326942DN1903_3)
- Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, 15(3), 118-121.
- Cardebat, D., Doyon, B., Puel, M., Goulet, P., & Joanette, Y. (1990). Formal and semantic lexical evocation in normal subjects. Performance and dynamics of production as a function of sex, age and educational level. *Acta neurologica belgica*, 90(4), 207-217.
- Cartwright, K. B., Marshall, T. R., Huemer, C. M., & Payne, J. B. (2019). Executive function in the classroom: Cognitive flexibility supports reading fluency for typical readers and teacher-identified low-achieving readers. *Research in developmental disabilities*, 88, 42-52.



- Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, 37(5), 715–730. <https://doi.org/10.1037/0012-1649.37.5.715>
- Cepeda, N. J., Cepeda, M. L., & Kramer, A. F. (2000). Task switching and attention deficit hyperactivity disorder. *Journal of abnormal child psychology*, 28, 213-226.
- Chelune, G. J., & Baer, R. A. (1986). Developmental norms for Wisconsin card sorting test. *Journal of Clinical and Experimental Neuropsychology*, 8, 219–228.
- Chevalier, N., Blaye, A., Dufau, S., & Lucenet, J. (2010). What visual information do children and adults consider while switching between tasks? Eye-tracking investigation of cognitive flexibility development. *Developmental Psychology*, 46(4), 955.
- Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children’s cognitive flexibility. *Developmental psychology*, 45(3), 782.
- Chevalier, N., Sheffield, T. D., Nelson, J. M., Clark, C. A., Wiebe, S. A., & Espy, K. A. (2012). Underpinnings of the costs of flexibility in preschool children: The roles of inhibition and working memory. *Developmental neuropsychology*, 37(2), 99-118.
- Clément, E. (2009). *La résolution de problème : À la découverte de la flexibilité cognitive*. Armand Colin.
- Clément, E. (2022). Successful Solution Discovery and Cognitive Flexibility. *Cognitive Flexibility: The Cornerstone of Learning*, 113-142.
- Cragg, L., & Chevalier, N. (2012). The processes underlying flexibility in childhood. *Quarterly Journal of Experimental Psychology* (2006), 65(2), 209-232. <https://doi.org/10.1080/17470210903204618>
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years : Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037-2078. <https://doi.org/10.1016/j.neuropsychologia.2006.02.006>
- Deák, G. O. (2000). The growth of flexible problem solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development*, 1, 157–191.
- Deák, G. (2003). The development of cognitive flexibility and language abilities. *Advances in child development and behavior*. [https://doi.org/10.1016/S0065-2407\(03\)31007-9](https://doi.org/10.1016/S0065-2407(03)31007-9)
- Deák, G. O., & Narasimham, G. (2014). Young children’s flexible use of semantic cues to word meanings : Converging evidence of individual and age differences. *Journal of Child Language*, 41(3), 511-542.

- Deák, G. O., & Wiseheart, M. (2015). Cognitive flexibility in young children : General or task-specific capacity? *Journal of Experimental Child Psychology*, 138, 31-53. <https://doi.org/10.1016/j.jecp.2015.04.003>
- Diamond, A. (2006). The Early Development of Executive Functions. In *Lifespan cognition : Mechanisms of change* (p. 70-95). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195169539.003.0006>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64, 135-168.
- Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, 18, 34-48. <https://doi.org/10.1016/j.dcn.2015.11.005>
- Drewe, E. A. (1974). The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. *Cortex*, 10(2), 159-170.
- Duvignau, K., Fossard, M., Gaume, B., Pimenta, M.-A., & Elie, J. (2007). Semantic approximations and flexibility in the dynamic construction and “deconstruction” of meaning. *Metaphor and Context: Linguagem em Discurso*, 7(3), 371–389.
- Ebersbach, M., & Hagedorn, H. (2011). The role of cognitive flexibility in the spatial representation of children’s drawings. *Journal of Cognition and Development*, 12(1), 32-55.
- Eslinger, P. J., & Grattan, L. M. (1993). Frontal lobe and frontal-striatal substrates for different forms of human cognitive flexibility. *Neuropsychologia*, 31(1), 17-28. [https://doi.org/10.1016/0028-3932\(93\)90077-D](https://doi.org/10.1016/0028-3932(93)90077-D)
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Stern, T. E. (2004). The Contribution of Executive Functions to Emergent Mathematic Skills in Preschool Children. *Developmental Neuropsychology*, 26(1), 465-486. [https://doi.org/10.1207/s15326942dn2601\\_6](https://doi.org/10.1207/s15326942dn2601_6)
- Ezekiel, F., Bosma, R., & Morton, J. B. (2013). Dimensional change card sort performance associated with age-related differences in functional connectivity of lateral prefrontal cortex. *Developmental cognitive neuroscience*, 5, 40-50.
- Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain: a journal of neurology*, 120(10), 1779-1791.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive development*, 10(4), 483-527.
- Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology*, 98(4), 611-625.

- Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404-411. <https://doi.org/10.1037/h0059831>
- Gros, H., & Gvozdic, K. (2022). Context, content effects and flexibility. *Cognitive Flexibility: The Cornerstone of Learning*, 203-232.
- Gros, H., Thibaut, J. P., & Sander, E. (2020). Semantic congruence in arithmetic: A new conceptual model for word problem solving. *Educational Psychologist*, 55(2), 69-87.
- Guilford, J.P. (1967). *The Nature of Human Intelligence*. McGraw-Hill, New York
- Hohl, K., & Dolcos, S. (2024). Measuring Cognitive Flexibility: A Brief Review of Neuropsychological, Self-Report, and Neuroscientific Approaches. *Frontiers in Human Neuroscience*, 18, 1331960.
- Iacono, E., Gros, H., & Clément, E. (2022). Training flexible categorization to improve arithmetic problem solving: A school-based intervention with 5th graders. In *CogSci 2022*.
- Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology*, 30(2), 190-200. <https://doi.org/10.1016/j.newideapsych.2011.11.001>
- Ionescu, T. (2017). The variability-stability-flexibility pattern: A possible key to understanding the flexibility of the human mind. *Review of General Psychology*, 21(2), 123-131.
- Jacques, S., Zelazo, P. D., Kirkham, N. Z., & Semcesen, T. K. (1999). Rule selection versus rule execution in preschoolers: an error-detection approach. *Developmental psychology*, 35(3), 770.
- Kavé, G., Kigel, S., & Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *Journal of clinical and experimental neuropsychology*, 30(3), 349-359.
- Kim, C., Johnson, N. F., Cilles, S. E., & Gold, B. T. (2011). Common and distinct mechanisms of cognitive flexibility in prefrontal cortex. *Journal of Neuroscience*, 31(13), 4771-4779.
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental science*, 6(5), 449-467.
- Kloo, D., Perner, J., Aichhorn, M., & Schmidhuber, N. (2010). Perspective taking and cognitive flexibility in the Dimensional Change Card Sorting (DCCS) task. *Cognitive Development*, 25(3), 208-217.
- Konishi, S., Nakajima, K., Uchida, I., Kameyama, M., Nakahara, K., Sekihara, K., & Miyashita, Y. (1998). Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nature neuroscience*, 1(1), 80-84.
- Korkman, M., Kirk, U., & Kemp, S. L. (2012). *Nepsy-II: manuel clinique et d'interprétation*. ECPA.

- Legare, C. H., Dale, M. T., Kim, S. Y., & Deák, G. O. (2018). Cultural variation in cognitive flexibility reveals diversity in the development of executive functions. *Scientific Reports*, 8(1), 16326. <https://doi.org/10.1038/s41598-018-34756-2>.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British journal of developmental psychology*, 21(1), 59-80.
- Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental science*, 12(1), 1-18.
- Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of neurology*, 9(1), 90-100.
- 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. <https://doi.org/10.1006/cogp.1999.0734>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134-140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 40(3), 255-265.
- Müller, U., Steven Dick, A., Gela, K., Overton, W. F., & Zelazo, P. D. (2006). The role of negative priming in preschoolers' flexible rule use on the dimensional change card sort task. *Child development*, 77(2), 395-412.
- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current directions in psychological science*, 21(2), 71-77.
- Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, 12(4), 313-324.
- Pan, X., & Yu, H. (2018). Different effects of cognitive shifting and intelligence on creativity. *The Journal of Creative Behavior*, 52(3), 212-225.
- Paulesu, E., Goldacre, B., Scifo, P., Cappa, S. F., Gilardi, M. C., Castiglioni, I., ... & Fazio, F. (1997). Functional heterogeneity of left inferior frontal cortex as revealed by fMRI. *Neuroreport*, 8(8), 2011-2016.
- Peng, A., Kirkham, N. Z., & Mareschal, D. (2018). Task switching costs in preschool children and adults. *Journal of experimental child psychology*, 172, 59-72.

- Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the dimensional change card sorting task?. *Infant and Child Development: An International Journal of Research and Practice*, 11(2), 93-105.
- Ravizza, S. M., & Carter, C. S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*, 46(12), 2924-2935.
- Rastelli, C., Greco, A., & Finocchiaro, C. (2020). Revealing the role of divergent thinking and fluid intelligence in children's semantic memory organization. *Journal of Intelligence*, 8(4), 43.
- Raynal, L.. (2022). Cognitive Flexibility and Analogy. *Cognitive Flexibility: The Cornerstone of Learning*, 175-202.
- Reiter-Palmon, R., Forthmann, B., & Barbot, B. (2019). Scoring divergent thinking tests: A review and systematic framework. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 144.
- Rosselli, M., & Ardila, A. (1993). Developmental norms for the Wisconsin Card Sorting Test in 5-to 12-year-old children. *Clinical Neuropsychologist*, 7(2), 145-154.
- Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, 55(1), 657-687. <https://doi.org/10.1146/annurev.psych.55.090902.141502>
- Scheibling-Sève, C., Gvozdic, K., Pasquinelli, E., & Sander, E. (2022). Enhancing cognitive flexibility through a training based on multiple categorization: Developing proportional reasoning in primary school. *Journal of Numerical Cognition*, 8(3), 443-472.
- Shu, B. C., Tien, A. Y., Lung, F. W., & Chang, Y. Y. (2000). Norms for the Wisconsin card sorting test in 6- to 11-year-old children in Taiwan. *The Clinical Neuropsychologist*, 14, 275–286.
- Tomer, R., Fisher, T., Giladi, N., & Aharon-Peretz, J. (2002). Dissociation between spontaneous and reactive flexibility in early Parkinson's disease. *Cognitive and Behavioral Neurology*, 15(2), 106-112.
- Titz, C., & Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychological research*, 78, 852-868.
- Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children : A meta-analytical study. *Learning and Individual Differences*, 23, 1-9. <https://doi.org/10.1016/j.lindif.2012.10.004>
- Yilmaz, F. N. (2019). Comparison of different estimation methods used in confirmatory factor analyses in non-normal data: A Monte Carlo study. *International Online Journal of Educational Sciences*, 11(4), 131-140..

- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS) : A method of assessing executive function in children. *Nature protocols*, 1(1), 297-301.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive development*, 11(1), 37-63.
- Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., & Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.
- Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control : II. The development of executive function in childhood. *Current Directions in Psychological Science*, 7(4), 121-126.