Bilingualism: Language and Cognition

cambridge.org/bil

Ready, steady, switch! Limited evidence for the role of executive functions in bilingual language control in children

Elisabet García González[1,2], Jussi Jylkkä[3] and Minna Lehtonen[1,2]

[1]Center for Multilingualism in Society Across the Lifespan (MultiLing), University of Oslo, Oslo, Norway; [2]Department of Psychology and Speech-Language Pathology, University of Turku, Turku, Finland and [3]Department of Psychology, Åbo Akademi University, Turku, Finland

Research Article

Cite this article: García González, E., Jylkkä, J. and Lehtonen, M. (2024). Ready, steady, switch! Limited evidence for the role of executive functions in bilingual language control in children. *Bilingualism: Language and Cognition*, 1–14
https://doi.org/10.1017/S1366728924000853

Received: 30 October 2023
Revised: 26 September 2024
Accepted: 26 September 2024

Keywords:
bilingualism; children; executive control; language control; language-switching; skill learning; task specificity

Corresponding author:
Elisabet García González;
Email: e.g.gonzalez@iln.uio.no

Abstract

We investigated the extent to which executive functions (EFs) are recruited in language switching in children in a cued picture-naming (CN) task. We expected to find associations between CN and EF tasks measuring inhibitory control and shifting. Another goal was to compare parent-reported children's everyday language control ability at home with their switching ability in the CN task and EF performance. The participants were mostly 5–7-year-old Norwegian–Spanish and Finnish–Swedish-speaking children (N = 45). The analysis was preregistered. Unexpectedly, the primary accuracy analysis showed positive associations between CN switching costs and EF performance in only one of the EF tests, flanker, and CN mixing costs were predicted only by the color-shape switch costs. Children's everyday language control ability did not show consistent significant associations with lab measures. Our study provides weak evidence for the view that EFs are engaged in language control when children have some years of bilingual experience.

Highlights

- Language switching is often assumed to engage executive functions (EFs).
- Children performed a cued naming switching task and EF tasks in the lab.
- Few significant associations were found between language switching performance and EFs.
- No consistent associations found between everyday language control ability and EFs.
- Results provide only weak support for the role of EFs in language switching in children.

1. Introduction

One of the most studied topics in the field of bilingualism in the last two decades is the bilingual cognitive advantage: the assumption that bilinguals can outperform monolinguals in executive functions (EFs) (Bialystok & Viswanathan, 2009) such as attention (Bialystok & Martin, 2004), working memory (Morales et al., 2013), inhibitory control (Hilchey & Klein, 2011), or cognitive flexibility, also known as task shifting (Prior & MacWhinney, 2010). The bilingual advantage hypothesis presupposes that (1) bilingual language use engages domain-general EFs, such as inhibitory control and shifting, and that (2) language switching can train EFs (Bialystok, 2009; Bialystok & Viswanathan, 2009). According to the first assumption, when bilingual speakers switch between languages, they must inhibit the language that was previously being used (thus engaging inhibitory control) and alternate languages in a similar manner to the way one would shift between non-verbal tasks. However, the bilingual advantage hypothesis has faced major challenges. A number of meta-analyses and systematic reviews have questioned whether a broad bilingual advantage exists (e.g., de Bruin et al., 2015; Donnelly, et al., 2019; Gunnerud et al., 2020; Lehtonen et al., 2018; Lowe et al., 2021; Monnier et al., 2022; Paap et al., 2013, 2015, 2018), thus challenging the training hypothesis of domain-general abilities and the idea that bilinguals could enjoy any such advantage. In addition, it is worth scrutinizing the more fundamental assumption as to whether language switching in fact engages EFs. If this is not the case, language switching could not train general EF, either. The present study addresses the assumption that language switching in children engages EFs.

The theoretical basis by which language switching engages EFs is that bilinguals' two languages remain active even when there is one target language for communication (Grosjean, 1989; Marian & Spivey, 2003). Thus, they must efficiently keep the languages from interfering with each other by exercising language control, which has been suggested to be regulated by domain-general executive control (Craik & Bialystok, 2006; Green, 1998; Green & Abutalebi, 2013), an assumption we call here the "domain-generality" account.

Some studies have indeed found significant links between tasks assessing domain-general EFs and language control in adult bilingual populations. For example, Declerck et al. (2017) found a positive association between similar tasks from the two domains, a non-verbal set-shifting task and a language-switching task, in a group of relatively highly proficient German-English bilinguals. Similarly, Graham and Lavric (2021) found associations between identical language switching and non-verbal switching paradigms in a group of highly proficient bilinguals. Furthermore, Jylkkä et al. (2018) found some associations between language switching and EF measures in high-proficiency bilinguals, albeit these were not replicated in a later study (Jylkkä et al., 2020). At the level of the brain, the results by de Baene et al. (2015), Declerck et al. (2021), and Lavric et al. (2019) suggested that neural circuits and ERP components related to domain-general EFs may be involved in language control as well.

On the other hand, studies also report findings not in line with the domain-generality account. For example, in two studies, Calabria et al. (2012, 2015) found no evidence in support of domain generality. Neither of these found significant correlations between a linguistic and a non-linguistic switching task in a group of highly proficient Spanish-Catalan bilinguals (2012) or different age groups of bilinguals (2015). The authors interpreted this result as evidence that the bilingual language control system is not fully dependent on the domain-general control system. These results were echoed in Branzi et al. (2016), and Magezi et al. (2012), as well as Wu and Struys (2021) who found dissociations between linguistic and non-linguistic EF tasks, thus challenging the hypothesis that language control in bilinguals is dependent on general executive control.

Some of the research that assumes domain generality has claimed that any potential executive demands in language switching are driven by the contexts in which speakers use their languages. The Adaptive Control hypothesis (ACH; Green & Abutalebi, 2013) proposes that the context of everyday language switching governs how strongly different EF processes are engaged since the speaker's linguistic context creates different cognitive demands upon them to use the right language. The ACH distinguishes three different language-switching contexts: a single-language, a dual-language, and a dense code-switching context. The single-language context applies when a speaker uses two languages in two clearly distinct contexts (e.g., at home versus at school). The dual-language context represents situations in which a speaker might be required to use one language with one speaker and switch to another with a different speaker within the same context, remaining aware of the interlocutor's needs. In the dense code-switching context, the speaker has relative freedom to use any language, based on which word or expression is more easily accessible, and to switch when desired without a break in communication, because the interlocutor is a bilingual speaker him- or herself. Under the ACH, the dual-language context is the most cognitively demanding because it involves almost all control processes proposed by Green and Abutalebi (2013): goal maintenance, conflict monitoring, interference suppression, salient cue detection, selective response inhibition, task disengagement, and task engagement. The plurality of control processes involved makes it more likely to engage EFs than the single-language context, which requires goal maintenance and interference control, or the dense context, which only requires opportunistic planning. Ultimately, the dual-language context demands more awareness of the linguistic environment, the interlocutor's needs, and assumed control to use the target language.

Recently, due to the mixed results, it has been questioned whether domain-general executive control is always engaged in language control in switching, even in dual-language contexts (Jylkkä et al., 2020; Lehtonen et al., 2023; Paap, 2018). Research in other areas of cognition has argued for the importance of automatization and addressed conditions where EFs are recruited in carrying out a particular task. In their Triarchic Theory of Learning, Chein and Schneider (2012) discuss three stages of learning from the acquisition of a new behaviour to its relative automaticity: the metacognitive system, the cognitive control network, and the representation system. According to the authors, these stages of learning support the establishment of new behavioural routines when presented with novel tasks, and ultimately provide a pathway for the automaticity of certain behaviours. While Chein and Schneider (2012) discuss how certain brain regions disengage earlier than others in the process of learning, they also raise the role of cognition in the learning of new tasks. In their proposal, EFs would likely be engaged in the learning of new tasks in novel contexts prior to the establishment of behavioural routines or skills (Chein & Schneider, 2012). Although these authors do not make claims about bilingualism, others have broadened this "skill learning account" into hypotheses of bilingual language control. Paap (2018) has applied these findings to the context of language learning by developing the "Controlled Dose Hypothesis." This account predicts that the process of language learning, similar to the process of learning a new task, may boost EF ability early on, but that such effects dissipate rapidly, as familiarity and automatization of the task increase. Therefore, this boost of EF ability is more likely to occur in the early stages of L2 learning. A similar account has been recently presented by Lehtonen et al. (2023; see also Jylkkä et al., 2017), who specify the assumptions that the skill learning view could take in the field of bilingualism. They propose that the reliance of language control on EFs is likely to diminish with accumulating bilingual experience. Hence, EFs would be more actively engaged when the subject is confronted with a novel task. They would make use of general inhibitory control or cognitive flexibility to resolve the newly presented task and to create strategies to perform it effectively. However, once the task is familiar enough, relatively automatic, task-specific processes have developed, after which the need to recruit EFs is smaller. Following this claim, bilingual speakers for whom switching is a daily activity – even in a dual-language context – might gradually recruit EFs to a lesser extent for this task.

As the amount of bilingual experience is a key factor in this skill-learning or task-specificity hypothesis, Lehtonen et al. (2023) entertain the assumption that children might show clearer associations between language control and EFs, with clearly less experience of language control than adults. We might thus presume that children have not yet accumulated enough experience in language switching for it to become an automatic, task-specific process. There is some evidence that language switching may rely on domain-general EFs in bilingual children (see Lehtonen et al., 2023, for a review), but there are thus far few studies directly addressing this relationship.

In understanding EFs in relation to language control, it is also important to consider how age and development affect these processes. Most of the hypotheses described above are rooted in adult brains and cognitive processes. However, cognition changes across our lifespan: from substantial development in the first few years of life, continuing to adulthood, and decline in old age. The prefrontal cortex, which matures greatly in the first 5 years of life (Best et al., 2009; Best & Miller, 2010), is highly interconnected with the development of EFs, as well as language. Research addressing

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

the developmental trajectories of children's cognition has found that children might initially rely on more reactive control mechanisms that later develop into more proactive control (Chatham, Frank, & Munakata, 2009; de Bruin et al., 2020). Considering continuous cognitive development, which can be influenced by different environmental factors, an interesting question is whether language switching engages domain-general EFs in children and whether such engagement might be seen more clearly in reactive than proactive control, as measured by different EF tasks.

To measure the processes involved in bilingual speakers' language control in dual-language-contexts, previous studies have often used cued picture-naming tasks (CN). In these tasks, participants are required to name pictures in two languages in two different types of blocks: a single-language block, and a mixed-language block. A visual cue (such as a flag) indicates the language that must be used for each trial. Alternating between languages creates a switch cost, which can be inferred by means of speakers' naming speed or accuracy when they switch languages. The cost is obtained through the difference in naming performance between switch trials and repetition trials in mixed blocks where both kinds of trials are present. Evidence has often shown that speakers are slower in naming a word in their L1 directly after naming a word in their L2 than they are in naming a word in their L2 after using their L1 (Meuter & Allport, 1999). This asymmetric switching cost has been taken as evidence of a need to inhibit the stronger language during the production of the weaker language, and that inhibitory control is necessary for language switching, although it has also been questioned whether switch costs are an index for reactive inhibition in bilingual language control (Bobb & Wodniecka, 2013; Gade et al., 2021). In addition to switching costs, CN tasks provide a measure for mixing costs, calculated as the difference between repetition trials in mixed blocks and single trials in single blocks, assumedly reflecting monitoring the use of languages or preparedness to switch.

According to the skill learning account, novel tasks are likely to rely more on EFs than familiar ones. It is therefore possible that a laboratory-based language switching task, which presumably entails some novel aspects compared to everyday language switching, shows higher correlations with EF task performance than switching in a natural environment. Here, we also ask whether the language control ability of children in an everyday dual-language context in the family is related to their EF performance. If everyday language switching ability relies on EFs in children, we should also see associations between children's everyday language control ability and their EF performance.

Thus, we study the connection between language-switching performance in the lab versus everyday language control ability in those children who are functioning in dual-language contexts in their homes. From a methodological point of view, there are few existing studies that compare lab task performance to everyday language use (for exceptions, see, e.g., Jylkkä et al., 2017; 2020). This is the case even though finding such associations would provide support for the ecological validity of commonly used lab tasks, such as the CN task (for criticism on this task, see, e.g., Blanco-Elorrieta & Pylkkänen, 2018).

1.1. Language switching studies in children

Very few studies have thus far explored language switching in children in a controlled lab environment with a CN paradigm (Gross & Kaushanskaya, 2015; Kubota et al., 2020; de Bruin et al., 2020). Kubota et al. (2020) investigated associations between CN and EF tasks. The authors explored whether development in executive control and bilingual experience predicted language control in 7–13-year-old bilingual "returnee" children. A CN task and a Simon task showed that cognitive development overall predicted language control, indicating an overlap between executive control and language control: CN mixing costs, assumedly indexing monitoring of language use, were predicted by exposure to the L2, and were modulated by improvement in performance in the Simon task. Their results seem to support the view that EFs are to some extent involved in language switching in children. Gross and Kaushanskaya's (2018) findings can also be taken to reflect the role of domain-general EFs in bilingual language control. They reported that nonverbal task-shifting abilities contributed to 5–7-year-old children's cross-language errors in the CN task as well as to naming speed in the non-dominant language.

Other studies have measured EFs in relation to language switching in children with either free play sessions (Kuzyk et al., 2019; Smolak et al., 2020; Kang & Lust, 2018; Gross & Kaushanskaya, 2020), or parental reports (Kaushanskaya & Crespo, 2019; Bosma & Blom, 2019). Most of these studies report that proficiency is an important factor that drives language switching in children, often arguing that language competence plays a more crucial role in children than EFs (Gross & Kaushanskaya, 2020; Smolak et al., 2020). There are, however, suggestions that domain-general EFs show associations with language switching in these non-laboratory settings as well (e.g. Gross & Kaushanskaya, 2020; Kuzyk et al., 2019), though not all studies or measures have shown this link (Kang & Lust, 2018; Kuzyk et al., 2019).

In their review, Lehtonen et al. (2023) summarize that while some evidence indicates a potential association between EFs and switching in children, the findings are not always consistent, although EF associations for mixing costs are more typically found in children than those for switching costs. The differences across ages and language experience of the participants in these studies suggest that EF-language switching associations, or lack thereof, could be compatible with both the domain-generality and the skill learning accounts. The authors highlight that, while these associations may be more likely in children than in experienced adults, it is at present unknown how much experience would be required to minimize these associations.

In sum, studies aiming to establish a direct connection between children's EF performance, language switching in the lab, and their linguistic home environment are rare. Given the limited research on these topics, we lack in-depth understanding as to whether language control in children, as measured in the lab, is associated with domain-general executive control, as well as with their everyday language control ability at home.

1.2. The present study

There is still limited evidence on the role of cognitive control in language switching in children, and how development affects these cognitive processes. The primary aim of our study was to understand the extent to which language switching engages EFs in children. According to previous findings (see Lehtonen et al., 2023, for a review), we expect to find associations between EFs and language control in the lab. In a dual-language context that the CN task represents, domain-general EFs should be engaged according to the ACH (Green & Abutalebi, 2013). Similarly, according to the skill learning or task-specificity framework, novel tasks should engage EFs, and we assume that young children might not have developed automatized, task-specific subroutines for language

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

switching yet (Lehtonen et al., 2023). Furthermore, we studied whether children's everyday switching ability is associated with their language control ability in the lab, and with EF performance. We assumed that if everyday bilingual language control engages EFs, we should also see associations between children's language control ability at home, as reported by parents, and EF performance. We also expected to see associations between the lab-based CN task and everyday language control ability.

To address these questions, bilingual children roughly aged 5–7 performed a CN language switching task and four EF tasks in the laboratory. Their everyday language control ability at home was probed by a parental questionnaire. The analyses of the study were preregistered.

2. Method

2.1. Participants

This study included parallel data collection of Norwegian–Spanish-speaking children in Norway and Finnish–Swedish-speaking children in Finland. Other than the language materials (i.e., language proficiency tests and the specific linguistic stimuli in the language-switching task and the color-shape task), the experiment was equivalent in both countries (see Section 2.2).

A total of 45 children (mean age, 76 months; SD, 0.53; range, 48–100; 23 boys) participated in this study: Nineteen Norwegian–Spanish speaking children were tested in Norway and 26 Finnish–Swedish children in Finland. We did not exclude children who were exposed to other languages as well – for instance, English was often the language of communication between the parents, especially in Norway – but we required that they were exposed to Norwegian and Spanish or Finnish and Swedish daily and that they were sufficiently fluent in each language pair to hold a conversation. To measure proficiency, we used Cross-linguistic Lexical Tasks (CLT) scores in Norwegian (Simonsen et al., 2014), Spanish (Cantú Sanchez, 2016), Finnish (Kunnari, 2013) and Swedish (Ringblom et al., 2014). We used the CLT scores as a proxy for dominance, which in most cases represented the language in which the child scored highest. In the rare cases when both scores were equivalent, we used the bilingual language experience calculator (BiLEC; Unsworth, 2013), which provides a measure for the cumulative length of exposure to each language, to determine language dominance. As Table 1 indicates, the Finnish–Swedish sample appears to be more dominant for Swedish, whereas the Norwegian–Spanish one is more balanced for each language. However, these children were highly proficient in all languages. We excluded children who had learning and cognitive difficulties such as ADHD, autism spectrum disorder, or hearing or visual impairments. We collected information on the SES of the parents in these families, the majority of which had at least one parent with a bachelor's degree (65%), a master's degree (35%), or a PhD (3%).

Table 1. Language proficiency and dominance information of each language for the Norwegian–Spanish and Finnish–Swedish participants (N = 45)

	Norwegian	Spanish	Finnish	Swedish
Proficiency % score	74%	70%	77%	87%
Dominance %	53%	47%	19%	81%

Note. Average percentage proficiency scores in the Cross-linguistic Lexical Tasks are reported. Average percentage dominance is obtained from the proficiency score and average exposure to each language reported in the BiLEC.

To assess the associations between everyday language control ability and language control ability in the lab, we narrowed the sample to a subset of families (n = 18) who reported a "stricter dual-language culture" which resembled a dual-language-like context at home. There each parent was relatively consistently using one language with the child and encouraging the child to use only that language when speaking to them. The motivation was that it is possible to obtain a measure of everyday language control ability only in families that enforce a dual-language context, in which there is a need to control the use of a particular language. We obtained this sample through the questions in the parental report that addressed (a) the extent to which the parent reported not mixing languages in a conversation with their child, (b) the degree to which each parent encouraged the child to respond in the language used, and (c) the extent to which the child was able to meet those demands. The response scale was from 0% to 100%. We selected the families with parents who responded 50% or more to question (b). Of those families, we selected the "stricter" parent and measured the child's switching behaviour (c) as a response to those demands. We also checked whether these parents reported that they themselves switched languages with their children (a), which did not seem to be the case for this group of parents. In the subset of the sample, at least one of the parents had a bachelor's degree (55%), a master's degree (33%), or a PhD degree (3%). Thus, the SES distribution for the strict culture sample, as compared to the full sample, was relatively similar.

2.2. Procedure and materials

In Norway, the data processing plan was assessed by the Norwegian Agency for Shared Services in Education and Research in order to ensure that the data collected in the project was processed in accordance with data protection legislation (reference number 408035). Data collection in Finland was evaluated and approved by the Ethics Committee for Human Sciences at the University of Turku.

Parents provided digital consent for their children to participate in this study. In Norway, children came to the laboratory for the two 1-hour-long sessions, while data collection in Finland took place in kindergartens, in a quiet room, in similar sessions. In either case, the time in between sessions was of 1 week at minimum but no further than 2 weeks apart. Sessions I and II were counterbalanced, such that half of the participants took Session II first. Session I consisted of an EF test battery of four cognitive tests as we wanted to include two tasks per EF domain, two inhibition and two set-shifting tasks: Flanker, which has been used to measure inhibition in children (e.g., Weintraub et al., 2013; Yang et al., 2011) and the Nonverbal Stroop Card Sorting Test (NSCST), often used to index inhibitory control at various ages from childhood to the ageing populations (Koch & Roid, 2012); the Dimensional Change Card Sorting test (DCCS), which is commonly used to assess task shifting in young children (Zelazo, 2006), and the color-shape task, designed to measure task switching (Rubin & Meiran, 2005). The order of the tests was counterbalanced with a Latin Square design. Session II included the CN task preceded by the CLT in the relevant languages, Norwegian and Spanish or Finnish and Swedish, to assess the child's comprehension and production of verbs in each language. The language assessment was always performed at the start of the session, and the languages of the CLT were counterbalanced. In the language switching task, children completed two single-language blocks, with order counterbalanced, followed by three mixed blocks.

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

Session I was carried out in the preferred language of the child by an experimenter fluent in that language, to make the child most comfortable in the testing situation. Session II was carried out by a bilingual experimenter, as the tasks included two languages. In addition to the two sessions the children participated in, parents filled out a survey about language-switching practices at home and responded to questions from the BiLEC (Unsworth, 2013) via telephone or Zoom, or filled out the questionnaire themselves if an appointment was not possible. All tasks were run in Presentation software (Version 0.80, Neurobehavioural Systems, Inc., Berkeley, CA, www.neurobs.com) on a laptop.

2.2.1. Picture-naming task

We designed a non-voluntary cued picture-naming (CN) task according to previous literature (e.g., Gross & Kaushanskaya, 2018; Kubota et al., 2020) to test children's language-switching abilities. The task consisted of five blocks: two single blocks and three mixed blocks, allowing us to provide several breaks for children to complete the tasks. Participants named aloud pictures in Norwegian/Spanish or Finnish/Swedish according to a given language cue. A total of 20 individual pictures were selected. The single blocks consisted of 20 experimental items, and in them, the pictures were to be named in only one language. In the mixed blocks, a cue informed the participant about which language to use for naming of the picture. In these blocks, the same pictures used in single blocks were repeated twice (40). For all blocks, the cues remained on the screen throughout each trial to reduce working memory demands. Practice blocks were administered prior to the single blocks and the first mixed block. There were three types of trials in the CN task. The single blocks consisted of single trials only, where the same language was repeated for all items. The mixed blocks included switch trials, where the target language was different from the previous trial, and repetition trials, for which the target language was the same as for the previous trial. We obtained two measures from these three trial types: (a) mixing costs, the difference in performance between single and repetition trials across the single and mixed blocks, and (b) switching costs, the difference in performance between switch and repetition trials within the mixed blocks.

The order of the trials was randomized in the single blocks and pseudorandomized in the mixed blocks. We created 8 lists for the mixed blocks to control for order effects, and to ensure a sufficient number of switches in each block (16; 8 for each language) and repetition trials (23). We chose a proportion of 40% switches to 60% repetition trials to avoid the predictability of the switches. There were no more than 4 consecutive trials of the same type. The total number of trials was 160; 40 single trials (20 in each language) and 120 mixed trials, with 48 switch trials and 69 repetition trials. The first trial of every mixed block did not count as either switch or repetition. The children completed 5 practice trials for each of the single-language blocks and 16 practice trials for the mixed-language block.

In contrast to studies on adults who often use flags as cues, we selected drawings of two different girls to make it more child-friendly. The participants were told that each girl, who also had a distinctly Spanish, Norwegian, Finnish, or Swedish name, could only speak one language (either Spanish or Norwegian in the experiment in Norway, and Finnish or Swedish in the experiment in Finland), and that the participants needed to make sure they would understand the words that were said to them. The oral responses were recorded for later analysis.

The pictures were selected from the MultiPic Project (Duñabeitia et al., 2018). The words were matched across languages (Norwegian with Spanish, and Finnish with Swedish) for mean frequency (*Ordforrådet*, Lind, et al., 2015 for Norwegian; *Spanish corpus* (*esTenTen*) for Spanish; the *Turun Sanomat* corpus for Finnish and the *Göteborgs-Posten* corpus for Swedish; Laine & Virtanen, 1999), age of acquisition when available (Stanford Wordbank, López Ornat, et al., 2005; Simonsen et al., 2014), and a number of alternative names. Cognates were avoided. Each picture appeared on a white screen in a speech bubble. The cue appeared slightly to the right and above the target picture. A cue was given in all blocks to help participants familiarize themselves with the girl who was supposed to speak each language. Each trial lasted for a maximum of 4 seconds. A trial began with a white screen and a fixation cross for 200 ms, followed by a blank screen for 500 ms, and then the picture appeared simultaneously with a visual cue (girl denoting language). Both the cue and the picture remained on the screen for 4000 ms, regardless of when the response was produced. There was a 500 ms interval between trials. The instructions for the single blocks were given in the corresponding language. For the mixed blocks, the instruction was given in both languages by a bilingual research assistant.

2.2.2. Executive tasks

Flanker task We used an adapted version of the flanker task (Eriksen & Eriksen, 1974) as a measure of inhibition and selective attention. In this version of the task, children were presented with pictures of 5 fish instead of arrows, based on Park et al. (2018). On congruent trials, all fish pointed in the same direction (e.g. >>>>>) and on incongruent trials, the central fish pointed in the opposite direction (>><>>). In neutral trials, a picture of vertical seaweed substituted all but the central fish. The child's task was to correctly identify the direction of the fish in the trials by using the left or right button in the response box. We focused on the flanker interference effect, which is the difference between the neutral and incongruent trials measured by means of accuracy rates or RTs. Each trial began with a fixation cross in the shape of a star on a blue background to simulate the sea. The maximum duration of each trial was 1700 ms with an inter-stimulus interval (ISI) of 1000 ms. There were 60 testing trials preceded by 18 training trials (6 congruent + 6 incongruent + 6 neutral). Children were required to respond accurately on at least 3 of the 6 training trials in order to proceed to the testing phase. In case they did not, 2 series of practice items were added. The testing phase consisted of 60 trials, including congruent (20), incongruent (20), and neutral (20) trials.

Nonverbal Stroop card sorting test An adapted, computerized version of the NSCST was used to assess children's inhibitory control. This is a widely used test for inhibitory control across populations from 3 to 70 years of age. The test is a non-verbal version of the classic Stroop test that is particularly suitable for young children (Koch et al., 2012). In this computerized version, the participant must place each card in one of the four stacks which each correspond to a specific color (red, blue, yellow, green). The cards are numbered in a specific order for all participants. The test consists of a congruent condition followed by an incongruent condition. The difference in performance between these two conditions is called the Stroop effect, which is used as a measure of inhibitory control. In the congruent condition, the cards include two rectangular shapes of the same color (e.g. blue) with a white cross in one of the colors. The participant must place the card in the right color. In the incongruent condition, each rectangular shape in

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

the card represents a different color (e.g. red and blue). The white cross in one of the colored rectangular shapes indicates where the card must be placed. The mean Stroop effect for each participant was used in the statistical analysis.

Dimensional change card sorting test (DCCS) The DCCS test is widely used to measure shifting in children from 3 to 7 years of age. We created and adapted a computerized version of the task similar to Park et al. (2018). In our version, the test consisted of two target cards (a blue fish and a red rabbit) that remained the same throughout the experiment. The task consists of a (pre-switch) color task and a (post-switch) shape task. Participants received a practice round of 6 items (3 red fish and 3 blue rabbits) in randomized order. Following the practice round, children performed the color task, where they received 3 red fish and 3 blue rabbits in randomized order. Children moved directly to the shape condition explained by the experimenter, where they again received 3 red fish and 3 blue rabbits that now had to be sorted according to their shape, and not their color. The task ended with a mixed cue condition, where a cue indicated whether children could respond according to color (rainbow) or shape (paw). In the mixed condition, the distribution was the following: 11 red fish (color) 11 blue rabbits (shape), 4 red fish (shape) and 4 blue rabbits (shape). In our study, we used average accuracy performance in the mixed block in the statistical analysis.

Color-shape task We designed a version of the color-shape task that emulated, as much as possible, the CN switching task with oral responses. As for the CN task, we also obtained switching and mixing costs from the different trial types. In contrast to the usual color-shape stimuli, we used cats and fish for the shape block to make it more child-friendly. We also selected red and blue for the color block. The cues that indicated color and shape were a rainbow and a paw (as in the DCCS). As in the CN task, the cues remained on the screen throughout each trial.

The task consisted of five blocks: a single (shape) block, a single (color) block, and three mixed (shape/color) blocks. There were 20 trials in the single block and 40 trials in the mixed blocks (following the 20 × 20 × 40 design of the picture-naming task). The trials appeared in randomized order in the single blocks and in pseudorandomized order in the mixed block. We created 8 lists to ensure the right number of switch trials (16 switches; 8 to color and 8 to shape in each mixed list) and repetition trials (23). The first trial of each block did not count as either switch or repetition. We assured there were no more than 4 consecutive trials of the same type. In addition, we counterbalanced the order of the stimuli (red/blue, cat/fish). Each trial had a maximum duration of 4 seconds. A trial started with a fixation cross presented for 200 ms, followed by a blank screen for 500 ms. Both the cue and target remained on the screen for a maximum duration of 4 seconds. A 500 ms inter-trial blank screen interval was presented before the onset of the following trial.

2.2.3. Language assessment tools: CLT

We used the CLT (cf. Section 2.1) to assess children's proficiency in the relevant language pairs. Each CLT includes 4 phases that assess the production and comprehension of verbs and nouns by means of object/action naming or pointing. To minimize the length of the session, children were tested on comprehension and production of verbs only. For comprehension tasks, children saw four action pictures on the screen, and they were instructed to point at the correct picture (e.g., Who is biting?). In the production part of the task, children were asked to name the action that appeared on the screen (e.g., What is she doing? – watching TV).

Responses were coded as correct or incorrect based on whether they pointed at or produced the target word. Responses were also written on the answer sheet for later analysis, to assess responses post-hoc, in case a child's untargeted response was due to his or her language variety (e.g., dialectal differences in Latin American versus European Spanish). The CLT scores were used to determine the child's dominant language, which was later used to explore asymmetric costs in the CN task.

2.2.4. Bilingual Language Experience Calculator

Parents completed the BiLEC (Unsworth, 2013) for language input and exposure via Zoom or telephone. This allowed us to obtain measures for the absolute and cumulative length of exposure to each language, as well as background and linguistic information about children, parents, and other members of the family who spend a significant amount of time with the child.

2.2.5. Language switching questionnaire

In addition to BiLEC, we asked parents to respond to a short survey on language practices about language mixing in the household. This allowed us to assess the extent of a home dual-language context and to select the participants who experienced a dual-language context in their family. See Section 2.1 Participants, for more information about this subsample.

3. Results

This analysis was preregistered in "As Predicted" prior to data analysis (08/18/2022; reference number 104957: https://aspredicted.org/j5q7v.pdf). All analyses were performed in R using mixed effects logistic regression models ((G)LMMs, package lme4, Bates et al., 2015). We focused on accuracy measures, in line with previous studies on children (Davidson et al., 2006; Diamond & Kirkham, 2005; Gross and Kaushanskaya, 2020), which argue that accuracy is a better index for performance in children than reaction times[1]. The dependent variable was always in a long format, while the predictors were included as means. Each of the models below was designed to answer one of four questions. Some additions were made to the initial pre-registration to facilitate model fit. Because these models are notoriously hard to fit and converge (Mundry, 2021), we z-transformed some of the covariates that were entered into the models, such as age and EF variables. These transformations did not impact the output of the models but generally facilitated the models to converge. The z-transformed variable for age in months was always included as an additional covariate in the models. Subject and Item were always added as random effects.

[1]Given that the tasks were computerized, we collected reaction times of the Flanker, the Color-Shape and the CN tasks. While primarily focusing on accuracy (see preregistration), we also ran three linear mixed effects models using reaction times following a comparable structure to that of the logistic regression models, but in this case the dependent variable was log-transformed reaction times from the CN task. The first model included the Flanker interference effect in reaction times in interaction with CN condition, and the second and third models included color-shape switch costs and color-shape mixing costs in reaction times in interaction with CN condition, together with age and trial number. Of these, the flanker interference effect and color-shape mixing costs were significantly positively associated with CN mixing costs. Tables and plots of these models can be found in the Appendix (see Figure A1 and Table A1).

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

Table 2. Mean scores and SD for all tasks and conditions

Measure	Accuracy (%)	SD (%)
CN single	93.78	8.11
CN repetition	85.40	22.30
CN switching	79.50	27.22
Flanker neutral	89.78	15.75
Flanker incongruent	83.54	22.70
Stroop congruent	95.71	25.73
Stroop incongruent	96.03	22.44
CS single	95.45	6.14
CS repetition	87.86	17.63
CS switch	82.29	19.14
DCCS color (pre-switch)	96.51	10.35
DCCS shape (post-switch)	81.01	24.86
DCCS mixed	67.50	25.11

Note. CN = Cued Naming task; CS = Color-Shape task. Means and SD are expressed in percentages.

Participants were excluded from the analysis if the child appeared distracted and responded randomly during the experiment.

In the following subsections we answer the four questions as formulated in the pre-registration: (1) Do children exhibit switching and mixing costs in CN? (2) Are switching and mixing costs in CN associated with EF performance in the laboratory? In addition, we investigate (3) To what extent is children's everyday language control ability associated with CN, and (4) with EF performance in the laboratory? In order to explore the latter two potential associations, we selected the families that reported creating a dual-language environment at home, that is, those families that enforced a stricter switching culture between parents and children, whereby children are encouraged to respond to each parent in one specific language. For this subsample of families, we asked whether there are associations between CN in the laboratory and children's everyday language control ability at home, as well as whether there are associations between EF performance in the lab and children's everyday language control ability at home. We summarize the descriptive statistics for all tasks in Table 2.

3.1 Switching and mixing costs in cued naming

The first model addressed whether children exhibited basic switching and mixing costs in CN in the laboratory, that is, whether children's performance was lower in switch trials compared to repetition trials, and in repetition trials compared to single trials. The model examined these with CN performance as the dependent variable, CN condition (repetition, switch, and single) as a fixed factor, and Subject and Item as random effects. This model revealed statistically significant switching ($E = -.47$, SD = .08, $z = -5.61$, $p < .001$) and mixing costs ($E = 1.05$, SD = .12, $z = 8.89$, $p < .001$) (Table 3). Participants were less accurate in switch trials, as compared to repetition trials of the mixed block, and in repetition trials as compared to single-block trials. When including age as a covariate in the model, we observed that accuracy improved with age for all the conditions (Figure 1a), but the effects remained also without the inclusion of this covariate.

Table 3. Estimates (standard error in parenthesis) for the switching and mixing cost model, where CN Condition is the main predictor, and for the asymmetry model, where Condition interacts with Language

Models	Predictor	CN switching cost	CN mixing cost
Switching and mixing cost	Condition	−0.47*** (0.08)	1.05*** (0.2)
Asymmetric cost	Language*Condition	0.02 (0.17)	−0.66* (0.2)

*:$p < .05$
***:$p < .001$.
Note. CN = Cued Naming task.

To explore a possible asymmetry in switching and mixing costs in CN, we ran a model including the interaction between language (L1/L2) and condition, and age as an additional covariate. Language, which was used here as a proxy for dominance, was determined by CLT production scores, as this domain was considered the most comparable measure to the CN task. For the purpose of this question, "L1" was assigned to the language with the highest score for the child, even though the large majority of children in the study were fairly proficient and balanced in the two languages. We nevertheless explored whether a difference in language proficiency – even if small – would influence the symmetry of switching and mixing costs. While we found the main effects of language and age, indicating that accuracy was higher in the assigned L1, and performance improved with age, the interaction between language and condition was not significant for switching costs ($E = .02$, SE = .17, $z = .14$, $p = .89$), but it was for mixing costs ($E = -.66$, SE = .26, $z = -2.57$, $p = .001$), which were larger in the L1 than in the L2[2] (Figure 1b).

3.2 Cued naming and EF tasks

For our second question regarding associations between the EF tasks and the language switching task, we created four models with CN accuracy as the dependent variable and the interaction of CN Condition with the mean of one EF measure variable at a time. Age was always added as an additional predictor, and it was significant for all the EF models. See Table 4 for a summary and Figure 2 for plots of these models.

The model examining the interaction between the CN Condition and the flanker task revealed a significant interaction for switching costs: the larger the flanker interference effect, the larger the switching cost ($E = .17$, SE = .08, $z = 2.09$, $p = .0036$) (Figure 2a). The flanker interference effect did not predict mixing costs ($E = -.10$, SE = .11, $z = -.11$, $p = .91$). In the model focusing on the interaction between the NSCST task and CN, the interaction was only marginally significant for mixing costs: the larger the Stroop effect, the larger the mixing cost ($E = .24$, SE = .14, $z = 1.72$, $p = .08$) (Figure 2b), and it was not significant for switching costs ($E = -.11$, SE = .08, $z = -.1.28$, $p = .19$). Thus, there was a tendency of inhibition tasks to be associated with the magnitude of switching and mixing costs in the CN task.

We ran two models to assess the associations between the color-shape task and CN: one model targeted the color-shape switching cost, and the other the color-shape mixing cost. The first model

[2]Deviating from the preregistration but based on previous studies (Gross & Kaushanskaya, 2015, 2018, 2020), we also tried including proficiency, measured by CLT production scores, as a fixed factor in additional analyses with CN and EF measures. While an additional covariate in the models explained some of the variation, the pattern of results did not change.

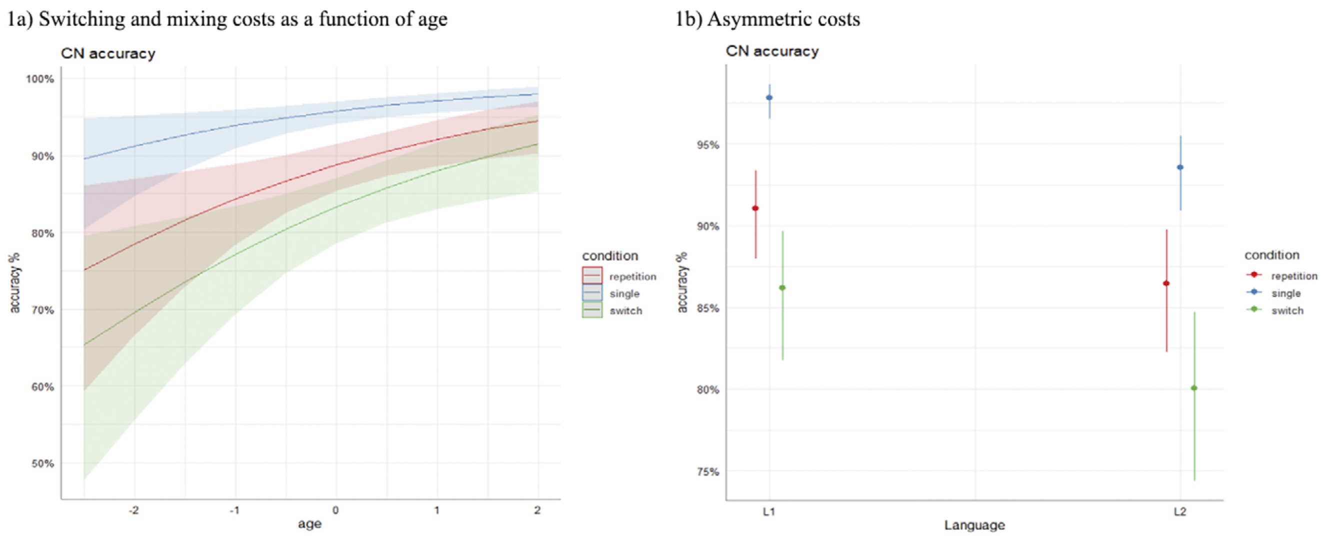

Figure 1. Models assessing basic switching and mixing costs and asymmetric costs for cued naming. (a) Switching and mixing costs as a function of age (b) Asymmetric costs.

Table 4. Estimates (standard error in parenthesis) for all EF interaction effects with CN condition, including non-significant ones in the CN task. Each line represents which EF measure each model analysed

Models	Predictor	CN switching cost	CN mixing cost
EF model 1	Flanker	0.17* (0.08)	−0.01 (0.12)
EF model 2	Stroop	−0.1 (0.08)	0.24 (0.14)
EF model 3	CS switching cost	−0.01 (0.01)	0.08*** (0.02)
EF model 4	CS mixing cost	0.04(0.01)	−0.004 (0.01)
EF model 5	DCCS	−0.08 (0.08)	−0.03 (0.12)

*:$p < .05$
***:$p < .001$.
Note. CN = Cued Naming task, CS = Color-Shape task.

revealed that the interaction between the switching cost of the color shape and CN condition was not significant for switching costs in CN ($E = -.01$, SE = .01, $z = -.71$, $p = .47$), but it was significant for mixing costs in CN ($E = .08$, SE = .02, $z = 3.92$, $p < .001$) (Figure 2c), suggesting an association between EF and the CN mixing cost. The larger the switching cost in the color-shape task, the larger the mixing cost in the CN task. The second model assessing the effect of mixing costs of the color-shape task on CN did not reveal a significant interaction for either switching or mixing costs in CN (Figure 2d).

Lastly, the model looking at the interaction between DCCS performance and CN was not significant for either switching ($E = -.08$, SE = .08, $z = -.77$, $p = .44$) or mixing costs of CN ($E = -.03$, SE = .12, $z = -.01$, $p = .99$). Age was the only significant predictor in the model. In addition to these models, we examined possible correlations between measures that tapped into the same component: Flanker and Stroop for inhibition, and color-shape and DCCS for shifting, but no statistically significant correlations were observed.

3.3 Cued naming and children's everyday language control

To analyze our third question regarding the associations between children's everyday language control ability as reported by the parents and CN in the lab, we created a model with CN accuracy as the dependent variable and CN condition in interaction with mean everyday language control ability (Child-switching) as fixed factors. Age was an additional predictor, and Subject and Item were added as random effects.

To assess associations between everyday language control ability and language control ability in the lab, we ran a model with the subset of families ($n = 18$) who reported a "stricter dual-language culture" (cf. Section 2.1). The variable "Child-switching" reflects the child's unwanted switching measured in percentages, in response to the parent who encourages the use of one language only. Therefore, an increase in this variable translates into a child's weaker ability to adhere to those demands.

We then analyzed the interaction of condition and child-switching tendency and its effect on CN accuracy in this subsample. It was not significant for either switching ($E < 0$, SE < 0, $z = .42$, $p = .67$) or mixing costs ($E < 0$, SE < 0, $z = .001$, $p = .99$) (Figure 3).

3.4 EF tasks and children's everyday language control

Our fourth question addressed associations between children's everyday language control ability, as reported by the parents, and EF. In this model, EF accuracy acted as the dependent variable (in long format) and the interaction of EF condition and mean everyday language control ability as fixed factors. Age was included as an additional covariate, and Subject and Item as random effects. We ran these models on the same subsample as for question 3.3. See Figure 4 for plots of the models. The first model did not reveal a significant interaction between the flanker interference and children's everyday language control ability ($E = .55$, SE = .37, $z = 1.52$, $p = .13$). The model on Stroop performance and everyday switching revealed a significant interaction in the unexpected direction ($E = .1$, SE = .42, $z = 2.31$, $p = .02$): the worse the children's everyday language control ability, and therefore more difficulties to stick to the demands of stricter dual-language context, the smaller the Stroop effect. The DCCS did not reveal a significant interaction ($E = .67$, SE = .28, $z = 2.49$, $p = .42$). The model exploring the interaction between child's language control ability and color-shape condition did not reveal significant interactions for either switching ($E = -.05$, SE = .18, $z = -.29$, $p = .77$) or mixing costs ($E = .42$, SE = .27, $z = 1.55$, $p = .12$).

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

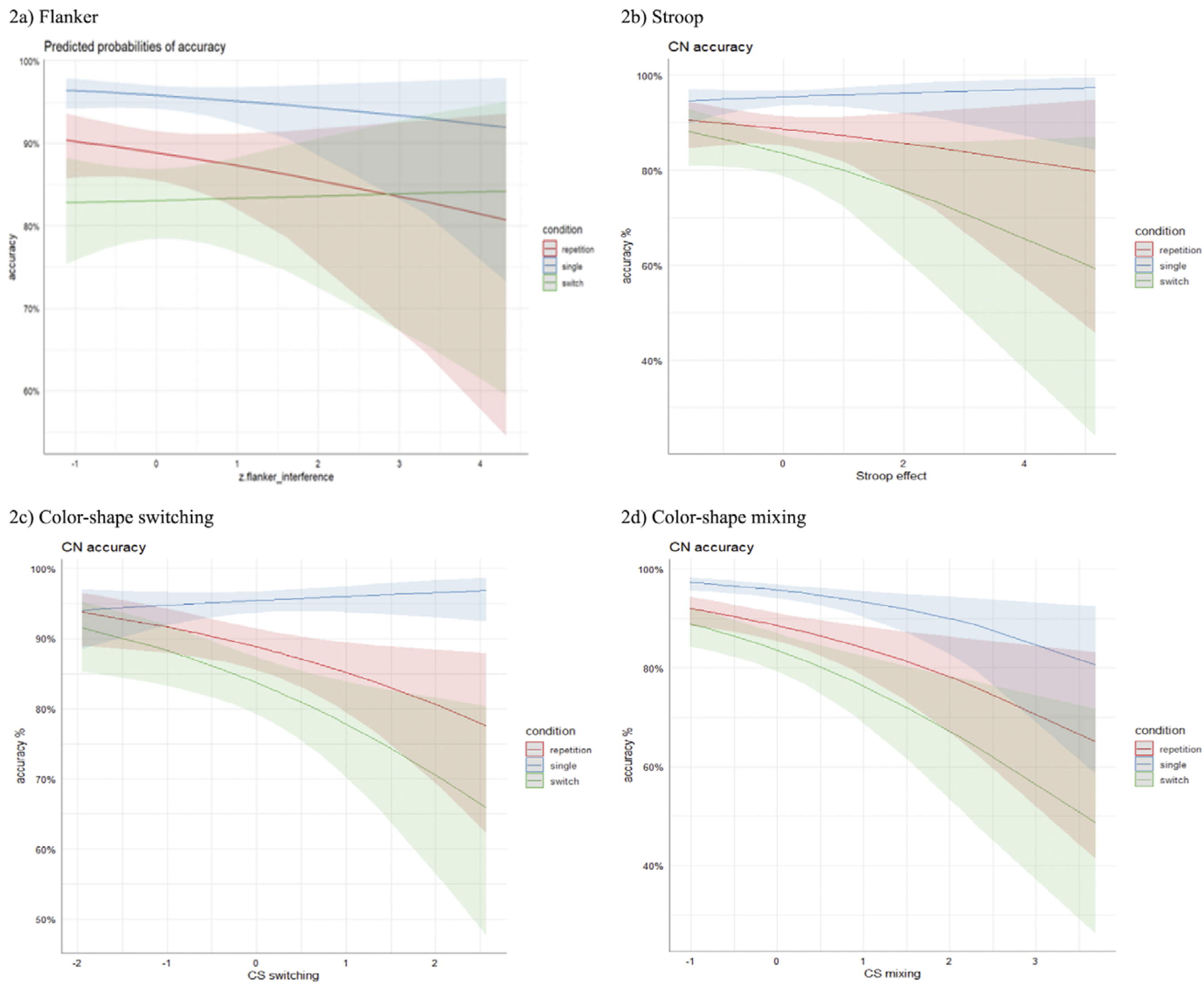

Figure 2. Models assessing the different (z-transformed) EF measures in interaction with the cued naming condition.

3.5 Summary of results

As predicted, the statistical analyses revealed overall switching and mixing costs in CN that decreased with age, but no evidence for asymmetric switching costs. However, we found asymmetric mixing costs: mixing costs were larger in L1 than in L2. The models exploring associations between EF performance and CN revealed some associations. Specifically, the flanker interference effect predicted switching costs in CN. Although we observed a trend for the Stroop effect to predict CN mixing costs, this effect did not reach statistical significance. We also found an association between the color-shape task and CN: switching costs in the color-shape task predicted mixing costs in CN; however, the mixing cost in the color-shape task was not a significant predictor for either switching or mixing costs in CN. We did not find any associations between the DCCS and CN.

Few direct associations were found between experience with a stricter dual-language culture at home and CN or EF performance in the lab. Children's ability to stick to the demands of a stricter language culture did not predict CN task performance in the laboratory. Among the five EF measures, we found only one significant association between EF performance and children's language control ability. Specifically, children's everyday language control ability at home predicted mixing costs in the Stroop task, with worse language control being associated with a smaller (better) Stroop effect.

4. Discussion

This study investigated the extent to which EFs are recruited in language switching in children in a controlled dual-language context using a CN paradigm. On the assumption that language switching engages EFs in children, our expectation was to find consistent associations between CN and EF performance. We also aimed to understand the extent to which children's everyday switching ability, as assessed by parental reports, is associated with EFs and language control ability in the lab.

Our analysis revealed significant switching costs, in line with previous research on switching in children (de Bruin et al., 2020; Gross & Kaushanskaya, 2015). We also found mixing costs, which are a frequent finding in studies of voluntary (e.g., de Bruin et al., 2020) and non-voluntary switching in children (e.g., Kubota et al.,

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

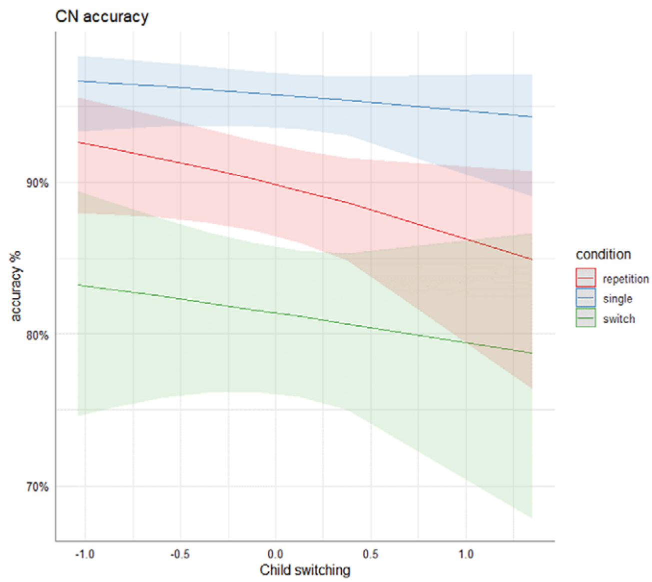

Figure 3. Model assessing children's everyday language control ability ("Child-switching") in interaction with the cued naming condition in the "strict dual-language culture" subset.

2020), as well as in cued switching in adults (e.g., Costa et al., 2006; Jylkkä et al., 2020; Meuter & Allport, 1999). Unlike Gross & Kaushanskaya (2020), we did not observe significant asymmetric costs in switching, although proportionally, children made more errors in their L2 (or less dominant language) than in their L1 (more dominant language) on average. However, we did find asymmetric mixing costs, similar to Kubota et al. (2020).

As to whether these switching and mixing costs are related to executive control, our data revealed only some associations between EF performance and CN measures. Specifically, of the two inhibition tasks, the flanker task predicted switching costs in CN, while the Stroop task only marginally predicted mixing costs in CN. The EF tasks tapping into shifting revealed an association between the color-shape task and mixing costs in CN, but no associations between the DCCS task and CN. Mixing costs in CN were positively associated with the switching cost in the color-shape: larger switching costs in the color-shape task predicted increased mixing costs in CN. Based on our data, therefore, EF performance was sometimes, though not consistently, associated with CN mixing, and only flanker performance was associated with CN switching.

Previous research suggests that young children initially rely on reactive control but gradually transition into proactive control

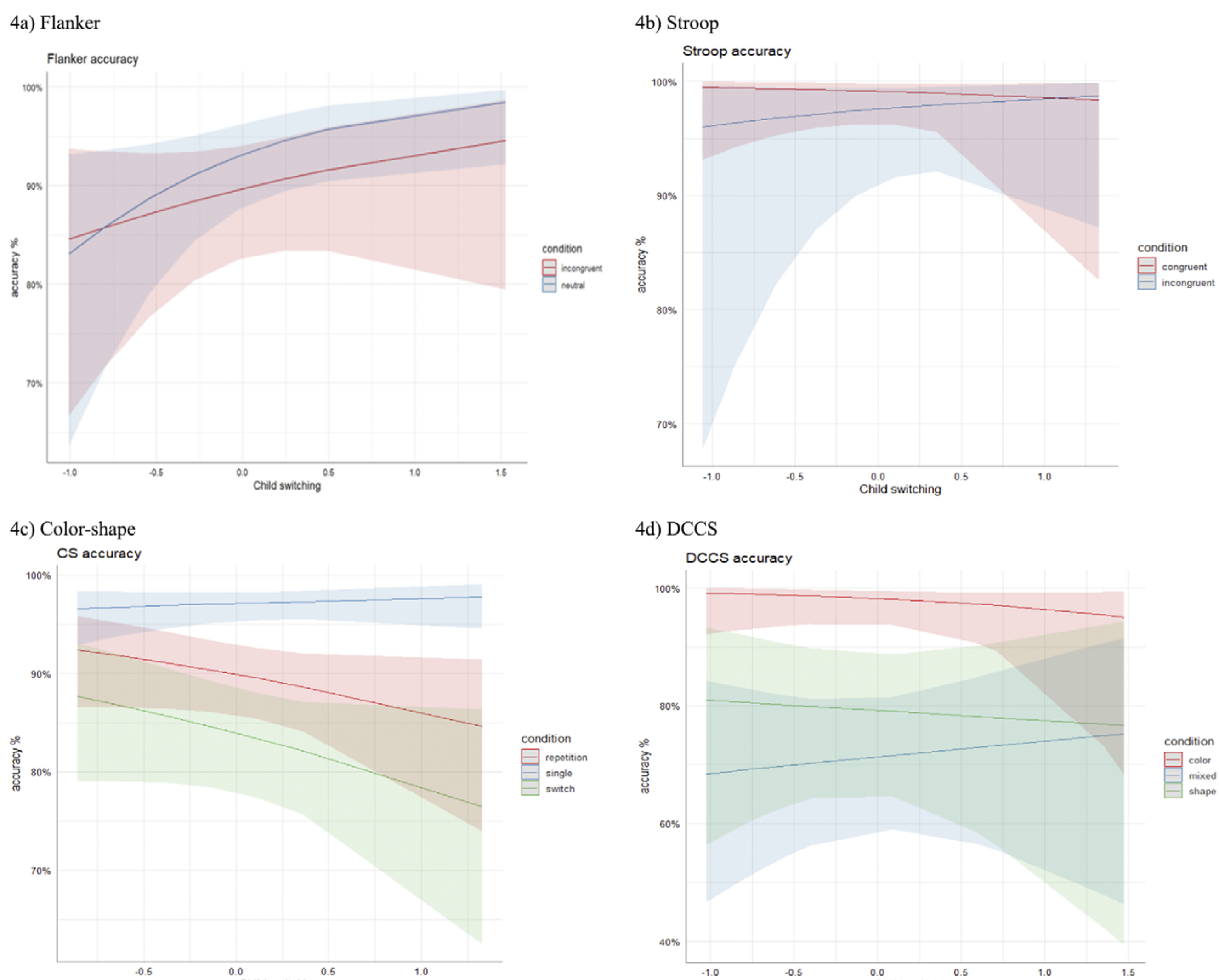

Figure 4. Models assessing the different (z-transformed) EF measures in interaction with children's everyday language control ability.

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

between 5 and 6 years of age (Gonthier et al., 2019). In the present study with mostly 5–7-year-old children, we observed that our measure of reactive control, the switching cost in the color-shape task, showed a significant positive association with the CN mixing costs in the primary analysis focusing on accuracy (see preregistration). In turn, our measure of proactive cognitive control, mixing costs in the color-shape task, was not associated with CN measures in accuracy. In the secondary response time analysis, however, this measure was positively associated with CN mixing costs. This latter finding could suggest that both reactive and proactive control processes had some role in language control in these children.

In general, observing significant positive associations between language mixing and switching costs could provide tentative evidence of the involvement of domain-general EFs in bilingual language control in children's language switching. This would be in line with the domain-generality framework (e.g., Green, 1998; Green & Abutalebi, 2013). Also, when considering this finding from the perspective of skill learning, Lehtonen et al. (2023) discuss that associations may be observed in children who might not have accumulated enough experience in language switching for task specificity to emerge. However, the majority of the analyses here did not show a positive association between EF measures and CN switching or mixing costs. Importantly, out of five EF measures, only one was positively associated with CN switching costs, and similarly, only one measure showed a significant association with CN mixing costs in accuracy. Thus, the evidence for EF involvement in bilingual language control was weak in the present study. When considering the secondary response time analysis (see Figure A1 in the Appendix), there was more evidence for the involvement of EFs in CN mixing costs: Both the flanker interference effect and color-shape (CS) mixing costs were positively associated with this measure. Interestingly, the recent skill-learning review (Lehtonen et al., 2023) predicts that language monitoring, assumedly reflected in the CN mixing costs, might be a domain where automatization is less likely to take place than in language switching, perhaps reflected in these significant CN mixing cost associations.

It should, however, be noted that out of eight analyses for language switching costs, only one was statistically significant, and out of eight analyses for language mixing costs, three were statistically significant (see Table A2 in the Appendix). In general, the fact that several separate models were run for the data may increase the risk of Type I error, potentially resulting in some false positives as well. Thus, overall, the majority of the present analyses did not show significant associations between children's language-switching performance and EFs. Our findings thus do not provide consistent support for the view that language switching engages domain-general EFs in children.

Language switching and mixing costs have often been interpreted to reflect the involvement of EFs, such as inhibition. In the present study, we set out to test whether this is the case in bilingual children. While we observed significant switching and mixing costs in the CN task, we did not find substantial evidence to support the view that these costs reflect domain-general control processes, in the kinds of bilinguals studied here. If we interpret the results from the skill-learning account's perspective, the lack of consistent language switching – EF associations might reflect a shift from domain-general to task-specific processes that have already taken place as a result of the bilingual experience that even these relatively young children have accumulated. With a familiar task such as language switching, domain-general EFs may not be recruited anymore.

In addition to addressing the degree of association between EFs and CN, our goal was to investigate the extent to which children's everyday language control ability relies on domain-general EFs. Moreover, we studied the associations between everyday language control ability and language control ability in the lab, which would be important for establishing ecological validity for this commonly used lab task. For these goals, we restricted our sample to those families who provided an environment where children's demands to keep the two languages separate were high enough to assume a dual-language context. For these families, we analyzed the children's ability to follow those demands, i.e., to use a fixed language with the parent, and the extent to which that translated into the CN task in the lab, which also simulates a dual-language context. We found no significant associations between these measures. In addition, we found only one statistically significant association between everyday language switching and EF performance, and that was in the unexpected direction. The results could thus suggest that everyday language switching does not rely on domain-general EFs, at least on those functions that the lab EF tasks are measuring.

An obvious limitation, however, is that the everyday language control analyses were performed on a subset of families (n = 18), which was considerably smaller than the complete sample. This might have constrained our statistical power and hindered the possibility of finding any significant associations between everyday switching versus CN and EF performance. Another potential question deals with the reliability and validity of parental reporting for measures of everyday switching ability and language culture in the home. These variables, obtained from parental reports, were central to understanding which families imposed dual-language context demands at home, and the extent to which children obeyed those demands, but they can be a noisy measure of language use and switching behaviours, and they can also be affected by parents' potentially negative attitudes towards language mixing between parents and children. However, it should also be noted that parental reports have been shown to be reliable in assessing expressive vocabulary (Dale, 1991) and general language skills in children (Garibaldi et al., 2021), and some other studies have used parental reports as the only proxy for language switching in children (Kaushanskaya & Crespo, 2019; Bosma & Blom, 2019). Another general limitation is children's performance in the Stroop task, which showed almost at-ceiling performance in our sample and may have repercussions for the results.

Reflecting on the current theoretical frameworks, our results as a whole only provide weak support for the domain-generality account that assumes that language switching engages domain-general EFs that are also used for other tasks. Although we found some associations between the EF tasks and CN, only two of them were statistically significant in our primary analyses. We thus conclude that these results do not show consistent associations between EFs and language switching in children. While a lack of associations cannot be taken as evidence to directly support any account, such results could, however, be compatible with the skill learning framework. This account predicts that EFs are not strongly involved in familiar tasks that can rely on increasingly automatic, task-specific subroutines stored in procedural memory. Even though it has been proposed that children, with their short bilingual experience, might not have developed such automatized subroutines yet (Lehtonen et al., 2023), the present results could suggest otherwise. The present participants were roughly 5–7-year-old children, most of whom had acquired the two languages simultaneously from birth, were overall relatively balanced in their language use, and thereby already had a few years of experience of bilingual language use.

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press

The specific characteristics of our bilinguals might explain our contrasting results with those of Kubota et al. (2020) who found an association between Simon's task and L2 CN performance. The participants in their sample were slightly older (7- to 13-year-olds) and had acquired their second language around age 5. Furthermore, these "returnee children", as described by the authors, had been tested upon return to their L1 environment after having spent between 2 and 4 years in their L2 environment. Given the differences in terms of age of acquisition and contexts in which the languages might be used by the two samples, it is not entirely surprising to find some divergence in the results of the two studies. In fact, these parallel results might provide an opportunity to reflect on how the age of acquisition and language experience might affect the relationship between EFs and language switching if we took skill learning as a starting point. In the present study, the two languages had been acquired and used simultaneously since birth, whereas the participants in the Kubota et al. study had fairly separate contexts for the two languages and assumedly less recent experience with mixing the two, perhaps making their language switching less automatic. According to skill learning, this relationship is not static but a dynamic one, and may change and adapt across the lifespan and be influenced by the speaker's individual experiences and changing linguistic contexts. As such, the skill learning view is compatible with accounts such as the ACH (Green & Abutalebi, 2013) by which the cognitive system is expected to adapt to the specific circumstances of the speaker's environment. As pointed out by Lehtonen et al. (2023), it is an open question as to how much and what kind of experience might be needed for automatization and task specificity to take place in bilingual tasks such as language switching.

5. Conclusions

Our preregistered study adds to a limited body of research on a potential connection between children's language switching and EFs, one fundamental assumption behind the bilingual executive advantage hypothesis. In contrast to this assumption, our results indicate quite a modest involvement of EF abilities in language-switching performance. In our primary accuracy analysis, only one of the five used EF measures significantly predicted mixing costs in the CN task, and, similarly, one of the measures predicted switching costs in CN. Furthermore, our study did not find a connection between children's everyday language control ability reported by the parents and CN or EF performance in the lab. Further research with a larger sample is needed to establish the replicability of these findings. Overall, the results offer only weak support for the view that domain-general EFs are involved in children's language switching. While the observed lack of associations does not provide direct evidence for any framework, the skill learning account could explain the findings by assuming that early balanced bilingual children have developed task-specific skills for language switching that no longer rely on domain-general EFs.

Supplementary material. To view supplementary material for this article, please visit http://doi.org/10.1017/S1366728924000853.

Acknowledgments. The authors thank Erika Ribu, Claudia Tarragó Pascual, and Mireia Blanch i Ferrer for their help in data collection in Norway, and Kajsa Lassila, Eveliina Viitala, Miia Hakala, Saara Söderholm, and Aino Mäntysaari for their help in stimulus preparation and data collection in Finland. We are grateful to the involved kindergartens for their collaboration. We also thank Teemu Laine and Klaus Wik for their help in programming the tasks, as well as Ingeborg Sophie Ribu for her valuable support in finalizing the experiment. Lastly, we thank Julien Mayor for his input at different stages of this study.

References

Bates, D., **Mächler, M.**, **Bolker, B.**, & **Walker, S.** (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Best, J. R., & **Miller, P. H.** (2010). A developmental perspective on executive function. *Child Development*, **81**(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x.

Best, J. R., **Miller, P. H.**, & **Jones, L. L.** (2009). Executive Functions after Age 5: Changes and Correlates. *Developmental Review: DR*, **29**(3), 180–200. https://doi.org/10.1016/j.dr.2009.05.002.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, **12**(1), 3–11. https://doi.org/10.1017/S1366728908003477.

Bialystok, E., & **Martin, M. M.** (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, **7**, 325–339.

Bialystok, E., & **Viswanathan, M.** (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, **112**(3), 494–500. https://doi.org/10.1016/j.cognition.2009.06.014.

Blanco-Elorrieta, E., & **Pylkkänen, L.** (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends in Cognitive Science*, **22**, 1117–1126. https://doi.org/10.1016/j.tics.2018.10.001.

Bobb, S. C., & **Wodniecka, Z.** (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology*, **25**(5), 568–585. https://doi.org/10.1080/20445911.2013.792822.

Bosma, E., & **Blom, E.** (2019). A code-switching asymmetry in bilingual children: Code-switching from Dutch to Frisian requires more cognitive control than code-switching from Frisian to Dutch. *International Journal of Bilingualism*, **23**(6), 1431–1447. https://doi.org/10.1177/1367006918798972.

Branzi, F. M., **Calabria, M.**, **Boscarino, M. L.**, & **Costa, A.** (2016). On the overlap between bilingual language control and domain-general executive control. *Acta Psychologica.*, **166**, 21–30. https://doi.org/10.1016/j.actpsy.2016.03.001.

Calabria, M., **Branzi, F. M.**, **Marne, P.**, **Hernández, M.**, & **Costa, A.** (2015). Age-related effects over bilingual language control and executive control. *Bilingualism: Language and Cognition*, **18**(65), 65–78.

Calabria, M., **Hernandez, M.**, **Branzi, F. M.**, & **Costa, A.** (2012). Qualitative differences between bilingual language control and executive control: evidence from task-switching. *Frontiers in Psychology.*, **2**, 399. 10.3389/fpsyg.2011.00399.

Cantú Sánchez, M. (2016). Cross-linguistic Lexical Tasks: Spanish version (CLT-ES). A part of LITMUS COST IS0804 Battery. Unpublished material.

Chatham, C. H., **Frank, M. J.**, & **Munakata, Y.** (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 5529–5533. http://doi.org/10.1073/pnas.0810002106.

Chein, J. M., & **Schneider, W.** (2012). The brain's learning and control architecture. *Current Directions in Psychological Science*, **21**(2), 78–84. https://doi.org/10.1177/0963721411434977.

Costa, A., **Santesteban, M.**, & **Ivanova, I.** (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **32**(5), 1057–1074. https://doi.org/10.1037/0278-7393.32.5.1057.

Craik, F. I. M., & **Bialystok, E.** (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences*, **10**(3), 131–138. https://doi.org/10.1016/j.tics.2006.01.007.

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research*, **34**(3), 565–571. https://doi.org/10.1044/jshr.3403.565.

Davidson, M. C., **Amso, D.**, **Anderson, L. C.**, **Diamond, A.** (2006). Development of cognitivecontrol 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

De Baene, W., **Duyck, W.**, **Brass, M.**, & **Carreiras, M.** (2015). Brain circuit for cognitive control is shared by task and language switching. *Journal of Cognitive Neuroscience*, **27**(9), 1752–1765. https://doi.org/10.1162/jocn_a_00817.

de Bruin, A., **Samuel, A. G.**, & **Duñabeitia, J. A.** (2020). Examining bilingual language switching across the lifespan in cued and voluntary switching contexts. *Journal of Experimental Psychology: Human Perception and Performance. Advance online publication*, **46**(8):759–788. http://doi.org/10.1037/xhp0000746.

de Bruin, A., **Treccani, B.**, & **Della Sala, S.** (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychological Science*, **26**, 99–107.

Decklerck, M., **Grainger, J.**, **Koch, I.**, & **Phillipp, A. M.** (2017). Is language control just a form of executive control? *Evidence for overlapping processes in language switching and task switching*. https://doi.org/10.1016/j.jml.2017.03.005

Declerck, M., **Meade, G.**, **Midgley, K. J.**, **Holcomb, P. J.**, **Roelofs, A.**, & **Emmorey, K.** (2021). On the connection between language control and executive control-an ERP study. *Neurobiology of Language (Cambridge*, **2** (Mass, 4), 628–646. https://doi.org/10.1162/nol_a_00032.

Diamond, A., **Kirkham, N.** (2005). Not Quite as Grown Up as We Like to Think: Parallels Between Cognition in Childhood and Adulthood. *Psychological Science*, **16**(4), 291297. https://doi.org/10.1111/j.0956.7976.2005.01530.x

Donnelly, S., **Brooks, P. J.**, & **Homer, B. D.** (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin and Review.*, **26**, 1122–1147.

Duñabeitia, J. A., **Crepaldi, D.**, **Meyer, A. S.**, **New, B.**, **Pliatsikas, C.**, **Smolka, E.**, & **Brysbaert, M.** (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, **71**(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261.

Eriksen, B. A., & **Eriksen, C. W.** (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**(1), 143–149. https://doi.org/10.3758/BF03203267.

Gade, M., **Declerck, M.**, **Philipp, A. M.**, **Rey-Mermet, A.**, & **Koch, I.** (2021). Assessing the evidence for asymmetrical switch costs and reversed language dominance effects – A meta-analysis. *Journal of Cognition*, **4**(1), 55. https://doi.org/10.5334/joc.186.

Garibaldi, A., **Venkatesh, L.**, **Bhat, J. S.**, & **Boominathan, P.** (2021). Relationship between parental report of language skills and children's performance among 3-year-olds: Implications for screening language among preschoolers. *International Journal of Pediatric Otorhinolaryngology*, **151**, 110943–110943. https://doi.org/10.1016/j.ijporl.2021.110943.

Gonthier, C., **Zira, M.**, **Colé, P.** & **Blaye, A.** (2019). Evidencing the developmental shift from reactive to proactive control in early childhood and its relationship to working memory. *Journal of Experimental Child Psychology*, **177**, 1–16, ISSN 0022-0965, https://doi.org/10.1016/j.jecp.2018.07.001.

Graham, B., & **Lavric, A.** (2021). Preparing to switch languages versus preparing to switch tasks: Which is more effective? *Journal of Experimental Psychology: General*, **150**(10), 1956–1973. https://doi.org/10.1037/xge0001027.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, **1**(2), 67–81.

Green, D. W., & **Abutalebi, J.** (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**, 515–530. https://doi.org/10.1080/20445911.2013.796377. PMID: 25077013; PMCID: PMC4095950.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, **36**(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5.

Gross, M., & **Kaushanskaya, M.** (2015). Voluntary language switching in English-Spanish bilingual children. *Journal of Cognitive Psychology*, **27**(8), 992–1013. https://doi.org/10.1080/20445911.2015.1074242.

Gross, M., & **Kaushanskaya, M.** (2018). Contributions of nonlinguistic task-shifting to language control in bilingual children. *Bilingualism: Language and Cognition*, **21**, 181–194. https://doi.org/10.1017/S1366.728916001097.

Gross, M., & **Kaushanskaya, M.** (2020). Cognitive and Linguistic Predictors of Language Control in Bilingual Children. *Frontiers in Psychology.*, **11**, 968. https://doi.org/10.3389/fpsyg.2020.00968.

Gunnerud, H. L., **Ten Braak, D.**, **Reikeras, E. K. L.**, **Donolato, E.**, & **Melby–Lervag, M.** (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *Psychological Bulletin.*, **146**, 1059–1083.

Hilchey, M. D., & **Klein, R. M.** (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, **18**, 625–658. https://doi.org/10.3758/s13423-011-0116-7.

Jylkkä J., **Laine M.**, **Lehtonen M.** (2020). Does language switching behavior rely on general executive functions? *Bilingualism: Language and Cognition*, **24**, 583–595. https://doi.org/10.1017/S1366728920000619

Jylkkä, J., **Lehtonen, M.**, **Lindholm, F.**, **Kuusakoski, A.**, & **Laine, M.** (2018). The relationship between general executive functions and bilingual switching and monitoring in language production. *Bilingualism: Language and Cognition*, **21**, 505–522. https://doi.org/10.1017/S1366728917000104.

Jylkkä, J., **Soveri, A.**, **Wahlström, J.**, **Lehtonen, M.**, **Rodríguez-Fornells, A.**, & **Laine, M.** (2017). Relationship between language switching experience and executive functions in bilinguals: an Internet-based study. *Journal of Cognitive Psychology*, **29**(4), 404–419. https://doi.org/10.1080/20445911.2017.1282489.

Kang, C., & **Lust, B.** (2018). Code-switching does not predict Executive Function performance in proficient bilingual children: Bilingualism does. *Bilingualism: Language and Cognition*, **22**, 366–382.

Kaushanskaya, M., & **Crespo, K.** (2019). Does Exposure to Code-Switching Influence Language Performance in Bilingual Children? *Child Development.*, **90**, 708–718. https://doi.org/10.1111/cdev.13235.

Koch, C., & **Roid, G.** (2012). A Non-verbal Stroop Task Assessment of Inhibition among Children with Autism. *Archives of Clinical Neuropsychology.*, **27**, 592–593.

Kubota, M., **Chevalier, N.**, & **Sorace, A.** (2020). How bilingual experience and executive control influence development in language control among bilingual children. *Developmental Science*, **23**, e12865. https://doi.org/10.1111/desc.12865.

Kunnari, S. (2013). Crosslinguistic Lexical Tasks: Finnish version (CLT-FI). A part of LITMUS COST IS0804 Battery. Unpublished material.

Kuzyk, O., **Friend M.**, **Severdija, V.**, **Zesiger, P.**, & **Poulin-Dubois, D.** (2019). Are there cognitive benefits of code-switching in bilingual children? A longitudinal study. *Bilingualism: Language and Cognition*, **23**(3), 542–553. https://doi.org/10.1017/S1366728918001207

Laine, M. & **Virtanen, P.** (1999). *WordMill Lexical Search Program*. Center for Cognitive Neuroscience. University of Turku.

Lavric, A., **Clapp, A.**, **East, A.**, **Elchlepp, H.**, & **Monsell, S.** (2019). Is preparing for a language switch like preparing for a task switch? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **45**(7), 1224–1233. https://doi-org.ezproxy.uio.no/10.1037/xlm0000636.

Lehtonen, M., **Fyndanis, V.**, & **Jylkkä, J.** (2023). The relationship between bilingual language use and executive functions. *Nature Reviews Psychology*, **2**, 360–373. https://doi.org/10.1038/s44159-023-00178-9.

Lehtonen, M., **Soveri, A.**, **Laine, A.**, **Järvenpää, J.**, **de Bruin, A.**, & **Antfolk, J.** (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, **144**(4), 394–425. https://doi.org/10.1037/bul0000142.

Lind, M., **Simonsen, H.G.**, **Hansen, P.**, **Holm, E.** & **Mevik, B.-H.** (2015). Norwegian words: A lexical database for clinicians and researchers. *Clinical Linguistics & Phonetics* **29**(4), 276–290. Downloaded DATE from: http://www.tekstlab.uio.no/ordforradet/en/

López Ornat, S., **Gallego, C.**, **Gallo, P.**, **Karousou, A.**, **Mariscal, S.**, & **Martínez, M.** (2005). *MacArthur: Inventario de desarrollo comunicativo. Manual y Cuadernillos*. Madrid: TEA Ediciones. ISBN: 84-7174- 820-7

Lowe, C. J., **Cho, I.**, **Goldsmith, S. F.**, & **Morton, J. B.** (2021). The bilingual advantage in children's executive functioning is not related to language status: A meta-analytic review. *Psychological Science*, **32**, 1115–1146.

Magezi, D. A., **Khateb, A.**, **Mouthon, M.**, **Spierer, L.**, & **Annoni, J.-M.** (2012). Cognitive control of language production in bilinguals involves a partly independent process within the domain-general cognitive control network:

Evidence from task-switching and electrical brain activity. *Brain and Language*, **122**(1), 55–63. https://doi.org/10.1016/j.bandl.2012.04.008

Marian, V., & **Spivey, M.** (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, **24**(2), 173–193. https://doi.org/10.1017/S0142716403000092.

Meuter, R. F., & **Allport, A.** (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, **40**(1), 25–40. https://doi.org/10.1006/jmla.1998.2602.

Monnier, C., **Boiché, J.**, **Armandon, P.**, **Baudoin, S.**, & **Bellocchi, S.** (2022). Is bilingualism associated with better working memory capacity? A meta-analysis. *International Journal of Bilingual Education and Bilingualism.*, **25**, 2229–2255.

Morales, J., **Calvo, A.**, & **Bialystok, E.** (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, **114**(2013), 187–202.

Mundry, R. (2021). *Linear Models in R [Workshop]*. Oslo, Norway: University of Oslo.

Paap, K. R. (2018). Bilingualism in Cognitive Science. In *The Cambridge handbook of bilingualism* (pp. 435–465). Cambridge: Cambridge UP.

Paap, K. R. & **Greenberg, Z. I.** (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*. **2013**(66), 232–258. [CrossRef] 22.

Paap, K. R., **Johnson, H. A.**, & **Sawi, O.** (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, **69**, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014.

Park, J., **Ellis Weismer, S.**, & **Kaushanskaya, M.** (2018). Changes in executive function over time in bilingual and monolingual school-aged children. *Developmental Psychology*, **54**(10), 1842–1853. https://doi.org/10.1037/dev0000562.

Presentation (2004) [(Version 0.80)Computer software]. Albany, CA: Neuro-Behavioral Systems. (www.neurobs.com)

Prior, A., & **MacWhinney, B.** (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, **13**, 253–262.

Ringblom, N., **Håkansson, G.**, & **Lindgren, J.** (2014). *Cross-linguistic Lexical Tasks: Swedish version (CLT-SE). A part of LITMUS COST IS0804 Battery*. Unpublished material.

Rubin, O., & **Meiran, N.** (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **31**(6), 1477–1491. https://doi.org/10.1037/0278-7393.31.6.1477.

Simonsen, H. G., **Kristoffersen, K. E.**, **Bleses, D.**, **Wehberg, S.**, & **Jørgensen, R. N.** (2014). The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*, **34**(1), 3–23. https://doi.org/10.1177/0142723713510997

Smolak, E., **de Anda, S.**, **Enriquez, B.**, **Poulin-Dubois, D.**, & **Friend, M.** (2020). Code-switching in young bilingual toddlers: A longitudinal, cross-language investigation. *Bilingualism: Language and Cognition*, **23**, 500–518. https://doi.org/10.1017/S1366728919000257.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, **16**, 86–110. http://doi.org/10.1017/S1366728912000284.

Weintraub, S., **Dikmen, S. S.**, **Heaton, R. K.**, **Tulsky, D. S.**, **Zelazo, P. D.**, **Bauer, P. J.**, **Carlozzi, N. E.**, **Slotkin, J.**, **Blitz, D.**, **Wallner-Allen, K.**, **Fox, N. A.**, **Beaumont, J. L.**, **Mungas, D.**, **Nowinski, C. J.**, **Richler, J.**, **Deocampo, J. A.**, **Anderson, J. E.**, **Manly, J. J.**, **Borosh, B.**, … **Gershon, R. C.** (2013). Cognition assessment using the NIH Toolbox. *Neurology*, **80** (11 Suppl 3), S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded.

Wu, R., & **Struys, E.** (2021). Language Dominance and Sociolinguistic Experience Are Related to Language Control and Domain-General Monitoring Control: An Investigation in Bilinguals Who Live in a Minority/Majority Sociolinguistic Setting. *Frontiers in Psychology*, **12**, 594–648. https://doi.org/10.3389/fpsyg.2021.594648.

Yang, S., **Yang, H.**, & **Lust, B.** (2011). Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism: Language and Cognition*, **14**(3), 412–422. https://doi.org/10.1017/S1366728910000611.

Zelazo, P. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, **1**, 297–301. https://doi.org/10.1038/nprot.2006.46.

https://doi.org/10.1017/S1366728924000853 Published online by Cambridge University Press