**The effect of metacognitive executive function training on children's executive function, proactive control, and academic skills**

Maki Kubota*[a], Lauren V. Hadley[b], Simone Schäffner[c], Tanja Könen[d], Julie-Anne Meaney[e], Candice C. Morey[f], Bonnie Auyeung[g], Yusuke Moriguchi[h] Julia Karbach[i] & Nicolas Chevalier[j]

## Author Note

[a] Department of Languages and Culture, UiT The Arctic University of Norway, Hansine Hansens veg 18, 9019, Tromsø, Norway, Email: maki.kubota@uit.no

[b] Hearing Sciences - Scottish Section, University of Nottingham, Glasgow Royal Infirmary, 10
16 Alexandra Parade, G31 2ER, Glasgow, Scotland, UK, Email: Lauren.Hadley1@nottingham.ac.uk

[c] Department of Psychology, University of Koblenz-Landau, Campus Landau Fortstraße 7, D-76829 Landau / Pfalz, Germany, Email: schaeffner@uni-landau.de

[d] Department of Psychology, University of Koblenz-Landau, Campus Landau Fortstraße 7, D-76829 Landau / Pfalz, Germany, Email: koenen@uni-landau.de

[e] Department of Psychology, University of Edinburgh, 7 George Square, EH8 9JZ, Edinburgh, Scotland, UK, Email: julieanne.meaney@ed.ac.uk

[f] School of Psychology, College of Biomedical and Life Sciences, Cardiff University, 70 Park Place, CF10 3AT, Cardiff, UK, Email: MoreyC@cardiff.ac.uk

[g] Department of Psychology, University of Edinburgh, 7 George Square, EH8 9JZ, Edinburgh, Scotland, UK, Email: Bonnie.Auyeung@ed.ac.uk

[h] Joetsu University of Education, School Education Yamayashikicho, 1 Joetsu, Niigata Prefecture, 943-8512, Japan, Email: moriguchi.yusuke.8s@kyoto-u.ac.jp

[i] Department of Psychology, University of Koblenz-Landau, Campus Landau Fortstraße 7, D-76829 Landau / Pfalz, Germany, Email: karbach@uni-landau.de

[j] Department of Psychology, University of Edinburgh, 7 George Square, EH8 9JZ, Edinburgh, Scotland, UK, Email: nicolas.chevalier@ed.ac.uk

**Funding information**

This work was supported by joint grants from the UK Economic and Social Research Council (ESRC, ES/N018877/1) and the German Research Foundation (DFG, KA 3216/2-1). Bonnie Auyeung was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.813546, the Baily Thomas Charitable Fund TRUST/VC/AC/SG/469207686, the Data Driven Innovation and the UK Economic and Social Research Council (ES/W001519/1) during the course of this work.

This study was not preregistered. All data have been made publicly available at the UK Data Service and can be accessed at
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=854956

## Abstract

The current study investigated the effects of metacognitive and executive function (EF) training on childhood EF (inhibition, working memory (WM), cognitive flexibility, proactive/reactive control) and academic skills (reading, reasoning, math) among children from disadvantaged backgrounds. Children (*N* = 134, *M* age = 8.70) were assigned randomly to the three training groups: (a) meta-cognitive training of basic EF processes (Meta-EF) (b) training of basic EF processes (Basic-EF) and (c) active controls (Active Control). They underwent 16 training sessions over the course of two months. No effects of EF and/or metacognitive training were found for academic outcomes. However, both Meta-EF and Basic-EF groups demonstrated greater gains than the Active Control group on proactive control engagement and WM, suggesting that EF training promotes a shift to more mature ways of engaging EF. Our findings suggest minimal near- and far-transfer effects of metacognitive training, but highlight that proactive engagement of EF can be promoted through EF training in children.



*Keywords*: Metacognitive training, executive function, SES, academic skills, cognitive development


*Public Significance Statements*: By demonstrating the potential effects of executive function training, with or without metacognitive reflection training for children from low socioeconomic backgrounds, we hope this work will help generate concrete educational instruments to counter socioeconomic disparities and improve children's opportunities for educational equality.

## 1. Introduction

As children grow older, greater need for autonomy calls for increasingly complex and flexible goal-directed behaviours. Such behaviours are supported by rapidly developing executive function (EF), the goal-directed regulation of one's own thoughts, actions, and emotions (Diamond, 2012) in early childhood (for a review, see Karbach & Unger, 2014). Young children are already expected to engage emerging EF on a daily basis, for instance, to follow teachers' instructions at school, take turns, raise a hand before talking, etc. Indeed, emerging EF in childhood predicts some indices of mental health, as well as professional success and wealth in adulthood (e.g., (Borella et al., 2010; Moffitt et al., 2011; Titz & Karbach, 2014). Most EF training studies have reported promising gains on untrained EF tasks (near-transfer effects; Kassai et al., 2019), albeit some studies finding no generalization of EF training to other cognitive skills (see Melby-Lervåg & Hulme, 2013 and Diamonf & Ling, 2016 for meta-analysis and systematic reviews). Findings have also been inconsistent in terms of transfer to untrained cognitive and academic skills (far-transfer effects; Diamond, 2012; Diamond & Lee, 2011; Karbach & Unger, 2014; Sala & Gobet, 2017; Titz & Karbach, 2014; see Strobach & Karbach, 2021 for a review). Our current study adopts an innovative approach by targeting metacognitive processes in addition to basic EF processes, to examine whether such combined training programs can lead to greater improvements in EF abilities, including changes in the way EF is engaged, and academic skills among children.

Metacognition, which predicts academic success and intelligence (Ohtani & Hisasaka, 2018), is a multidimensional construct that encompasses both *metacognitive knowledge*, which refers to knowledge about oneself, the task, and strategies, and *procedural metacognition*, which entails identification and assessment of one's learning, performance, and strategy use (Brown, 2017; Roebers, 2017; Veenman et al., 2006). Following Nelson and Narens (1990), it is generally accepted that two aspects of procedural metacognition operate in concert. *Metacognitive monitoring* allows self-reflection on and self-evaluation of ongoing cognitive activity, including performance, progress towards goal attainment, and changes in task demands, while *metacognitive control* uses this information for online adjustment of ongoing activity (e.g., Bryce & Whitebread, 2012; Lyons & Zelazo, 2011; Roebers, 2017). Both aspects of metacognition develop with age and schooling during childhood (e.g., Schneider, 2008), although age-related progress in metacognitive control may lag behind metacognitive monitoring (e.g. Krebs & Roebers, 2010; Bryce et al., 2015). For instance, 9-year-olds, but not 6-year-olds, spend more time studying difficult than easy items in a memory task (i.e., greater metacognitive control in older children), even though younger children can already tell which items are harder to learn (i.e., they already show efficient metacognitive monitoring; Destan et al., 2014).

Several frameworks tightly link metacognition and EF at the conceptual level (Bryce et al., 2015; Kälin & Roebers, 2022; Roebers, 2017; Roebers & Feurer, 2016). For instance, age-related gains in monitoring and/or self-reflection abilities have been proposed to drive the development of both metacognition and EF (e.g., Roebers, 2017; Lyons & Zelazo, 2011). EF has also been regarded as a main contributor to metacognition development (e.g., one may more flexibly select a suitable strategy when efficient EF skills allow successful implementation of a wide array of strategies), which has received empirical support (Bryce et al., 2015; Kalin & Roebers, 2022; Roebers et al., 2012; but see Spiess et al., 2016). Conversely, EF development has been hypothesised to be driven in part by metacognitive gains (e.g., Chevalier, 2015; Zelazo, 2015). Accordingly, EF development is not so much driven by an quantitative increase in control resources, but by more flexible engagement of these resources as a function of changing task demands with age. Indeed, children initially engage EF in a rigid, undifferentiated, and suboptimal fashion independent of specific task demands, but show increasingly flexible EF engagement with age, as suggested by differentiation of the processes engaged to support inhibition, working memory, and cognitive flexibility (e.g., Karr et al., 2018; Lee et al., 2013), as well as emergence of new control strategies (e.g., verbal strategies; Cragg & Nation, 2010; Fatzer & Roebers, 2012). Greater flexibility in EF engagement may rely at least in part on growing metacognitive control and monitoring, including better error detection and feedback processing with age (e.g., Andersen et al., 2014; Chevalier et al., 2010; DuPuis et al., 2015; Hadley et al., 2020), in order to better represent variations in task demands and adjust *how* and *how much* control is engaged accordingly. For instance, when given the choice between tasks requirring more or less cognitive effort/EF engagement, 11-12-year-olds and adults, but not 6-7-year-olds, preferentially select the task that requires less effort, suggesting greater sensitivity to and/or use of variations in task demands to minimize EF engagement with age (Niebaum et al., 2019; see also Ganasan et al., 2021; O'Leary & Sloutsky, 2017).

Metacognitive development may contribute in particular to increasingly proactive engagement of EF processes during childhood—a major aspect of EF development (Munakata et al., 2012) that predicts academic skills over and beyond working memory, inhibition, and set-shifting performance (Kubota et al., 2020; Skau et al., 2022; Wang et al., 2021). Specifically, young children tend to engage EF reactively, down or up-adjusting EF engagement as changes in task demands occur (e.g., figuring out directions while already walking). In contrast, older children more flexibly engage EF either reactively or proactively, anticipating and actively preparing for upcoming tasks (e.g., looking up walking directions before going to a new place), depending on specific task demands (e.g., Blackwell et al., 2014; Chatham et al., 2009; Chevalier et al., 2015; Lucenet & Blaye, 2014; Voigt et al., 2014). Recent findings suggest that metacognitive development may contribute to more flexible engagement of reactive or proactive control with age, by allowing children to monitor how they engage EF (reactively or proactively), how well they did it, and whether it was successful given current task demands. Consistently, when given as long as they want to proactively prepare for an upcoming task, 6-year-olds were less likely than 10-year-olds to start the task before being fully prepared (even though they performed better after full than partial preparation), suggesting an age-related increase in metacognitive monitoring of proactive EF engagement (Chevalier & Blaye, 2016). Furthermore, 5- and 6-year-olds can engage control more proactively, when reactive control is made more difficult (Chevalier et al., 2015) or they are encouraged to monitor their performance (Hadley et al., 2020), suggesting that performance monitoring may be key to engagement of proactive control in younger children.

If metacognition contributes to EF development, including greater proactive control engagement, a critical question is whether training metacognition can improve EF, the same way it benefits mathematical and reading abilities in children (Schaeffner et al., 2021). Recent promising findings suggest that metacognitive training itself may enhance EF performance in clinical populations (Kajka, 2019, Haugen et al., 2023). For instance, Kajka (2019) examined the effect of metacognitive training on working memory in children with ADHD and Haugen et al. (2023) found Goal Maintenance Training effects on self-reported executive functioning for psychosis participants. However, evidence is missing in typically developing children. In fact, the few studies investigating mindfulness training support this view. Mindfulness training shares some characteristics with metacognitive training, because participants are instructed to reflect on current actions or thoughts and their context clearly and objectively (Shapiro et al., 2015). The training exercises are usually not only cognitive but include strategies like breathing exercises, sitting mediations, and more movement-based relaxation activities. Indeed, mindfulness training improved attention, concentration, and social-emotional abilities in children (Flook et al., 2010; Schonert-Reichl & Lawlor, 2010; Zelazo et al., 2018). However, these findings are based on the reports of teachers and parents, and therefore evidence for effects on behavioural measures or educational outcomes (such as the tests for EF and achievement in the current study) is still missing.

Combining metacognition and EF training may be especially powerful to enhance EF and in particular, flexible engagement of EF processes (Marulis et al., 2020). To our knowledge, only a handful of studies have combined EF and metacognitive training. Some work has shown metacognitive strategy training can enhance children's working memory and reading comprehension (Carretti et al., 2017; Partanen et al., 2015). However, these results were confounded by several factors such as the lack of a training group that only received process-based EF training (Carretti et al., 2017) or the lack of control activities for the basic-EF training only group (Partanen et al., 2015). Following this, two studies have compared EF and metacognitive training (Meta-EF) to EF training alone (Basic-EF) and to active controls (Control) in order to examine the scaffolding effect of metacognitive training on EF training in either 5-year-olds (Pozuelos et al., 2019) or 9- to 14-yeard-olds (Jones et al., 2020). Preschoolers in Pozuelos et al.'s study showed changes in neural markers for conflict processing at immediate post-test for the Meta-EF group only (in comparison to Basic-EF and Control group), although no group differences were observed at the behavioural level. In addition, they also found significant improvements in fluid intelligence for the Meta-EF group. Jones et al.'s study, on the other hand, found basic EF training (regardless of whether it was combined with metacognitive training) to contribute to greater working memory and mathematical performance observed at immediate post-test in 9- to 14-year-olds. However, working memory gains were better maintained for the Meta-EF group than the Basic-EF group three months after training. In sum, training metacogntive reflection alongside EF may translate into both near- (i.e., training effects of targeted EF processes) and far-transfer effects (i.e., training effects of untrained academic skills) by combining a strategy and process-based approach to training. However, existing studies have focused on only one EF and/or a restricted age range, potentially explaining why no behavioural effect on EF performance was observed at immediate post-test and limiting the generalizabity of the beneficial effect of combined metacognition and EF training. Importantly, none of these studies have examined the key question of whether metacognitive training leads to more mature ways of engaging EF such proactive control engagement.

The present study investigated whether combined metacognition and EF training in the context of not just one but all three main EF domains (inhibition, working memory, and shifting) may lead to immediate behavioural gains in terms of EF performance and far transfer to other cognitive and academic skills, relative to both EF training alone and active controls. Critically, it examined whether metacognition+EF training can elicit a shift to more proactive engagement of EF. To date, it is unknown whether proactive control can be enhanced through EF training and whether adding metacognition training yields additional gain. Yet, this is an important question given theoretical proposals that metacognition may play a key role in increasingly flexible engagement of reactive and proactive control as a function of task demands with age.

The training programme targeted mostly metacognitive monitoring (e.g., reflection on task demands, strategy use, and performance), given theoretical accounts that monitoring may be key to both metacognition and EF (Chevalier, 2015; Lyons & Zelazo, 2011; Roebers, 2017; Zelazo, 2017), but also included metacognitive control training (e.g., generation and selection of strategies) to maximise the potential effects of metacognitive+EF training. Given that transfer effects in previous studies were larger after working memory and shifting training than after inhibition training (Karbach & Unger, 2014; Thorell et al., 2009), we assessed the effects of training on each domain of EF separately by administering trainings which engaged in working memory, inhibition, and shifting. We specifically targeted low SES children (although our study does not look into SES effects of training gains), because process-based trainings targeting EF often resulted in compensation effects (i.e., the largest benefits in participants with the lowest cognitive baseline abilities; e.g., (Karbach et al., 2015, 2017). Thus, targeting this specific population may most likely induce greater transfer, compared to focusing on higher SES children. We expected both EF and metacognition+EF training to elicit near-transfer effects on working memory, inhibition, and shifting, relative to the controls. We expected these effects to be greater after metacognition+EF training than EF training alone. Critically, we predicted that the Meta-EF group would shift to more proactive EF engagement and show greater gains in academic skills (math; Jones et al., 2020 and/or matrix reasoning; Pozuelos et al., 2019), to a greater extent than the other groups.

## 2. Method

### 2.1 Participants

The participants in the study were 155 children (Mean age = 8.73 years, *SD* = .72; N female = 82) recruited from schools in Scotland (N = 93) and Germany (N = 62). In order to determine the deprived areas in each country, we used the Scottish index of multiple deprivation (SIMD) to identify the most deprived areas in Scotland (lowest 10%), then mapped each school's catchment area to the SIMD map to identify eligible schools. In Germany, most of the children (75%) were recruited in areas with a regional at-risk-of-poverty rate above the national mean (Federal Statistical Office, 2019). Ten participants who withdrew during the study and 11 participants who reported to have a diagnosis of developmental disorder, impaired hearing and vision, and/or acquired brain injury were excluded from the data analysis, leaving the final pool with 134 participants (Mean age = 8.70 years, *SD* = .97; N female = 71). The ethics boards at the [xxx] approved the study. This study was not preregistered. Parental consent was received for all participants and children received small presents such as stickers, stationaries, and books as compensation for participating in the study. All data have been made publicly available at the UK Data Service and can be accessed at [xxx].

### 2.2 Procedures

The participants in each country were randomly allocated to one of three different training groups: (a) MetaEF (b) BasicEF (c) Control. The statistical description of the three groups is indicated in Table 1. The groups did not differ on age, vocabulary, or processing speed ($p$'s > .31).

Table 1. Statistical description of the three training groups

| Group | N | Mean Age (*SD*) | Vocabulary (*SD*) | Processing Speed (*SD*) |
|---|---|---|---|---|
| MetaEF | 46 | 8.68 (1.04) | 25.34 (3.90) | 30.74 (8.63) |
| BasicEF | 42 | 8.66 (0.98) | 24.42 (4.67) | 28.10 (6.42) |
| Control | 46 | 8.75 (.93) | 25.33 (3.93) | 30.95 (6.77) |

### 2.3 Procedure

The study comprised two pre-test sessions (1 hour each), 16 training sessions (45 minutes each), and two post-test sessions (1 hour each) as indicated in Figure 1. Training commenced within two weeks of the pre-test sessions and was conducted two to three times per week for every participant. The post-test sessions were administered within two weeks of the final training session. In the first pre-test session, the participants completed the matrix reasoning and vocabulary subtests of the Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2004), as well as a mathematical reasoning task (WIAT-III; Wechsler, 2017), reading comprehension task (WIAT-III; Wechsler, 2017), and a processing speed task (colour-naming test). They also completed a range of questionnaires assessing metacognition, mindset, motivation, and personality and wellbeing. In the second pre-test session, they completed three executive function tasks (Backwards Corsi, Anti-saccade, Cued Task

Switching, see below). The same procedure was followed for the two post-test sessions (excluding the vocabulary and processing speed tasks, as we did not expect training to affect these abilities).

Training session 1 introduced the three primary executive function tasks (N-back, ARTS, AX-CPT, described in detail below) that were used in the first 13 training sessions. These following twelve sessions each focused on two of these tasks, rotating through all three as indicated in Figure 1. In each of sessions 14-16, a novel EF task was presented, in order to facilitate generalisation of training to new contexts. These tasks were all embedded and linked within a cover story that involved two elves, Lessa and Leo, who live in a magic kingdom called Asfallon. The children were asked to go on an adventure with them and help them collect magic points through completing different games that involved magicians, dragons, trolls, and other magic creatures. Sessions 2-16 followed a standard structure alternating between solo work on EF tasks (approximately 30 minutes), and pair work on metacognitive (for the Meta-EF group) or creative tasks (for Basic-EF and Control groups). The pair work lasted for approximately 15 minutes, split into 5-minute exercises at three points during each session. The set of stimuli used for EF training tasks changed after session 8 to facilitate generalization of the newly acquired skills across different materials. The sessions were always supervised by one to two members of the research team.

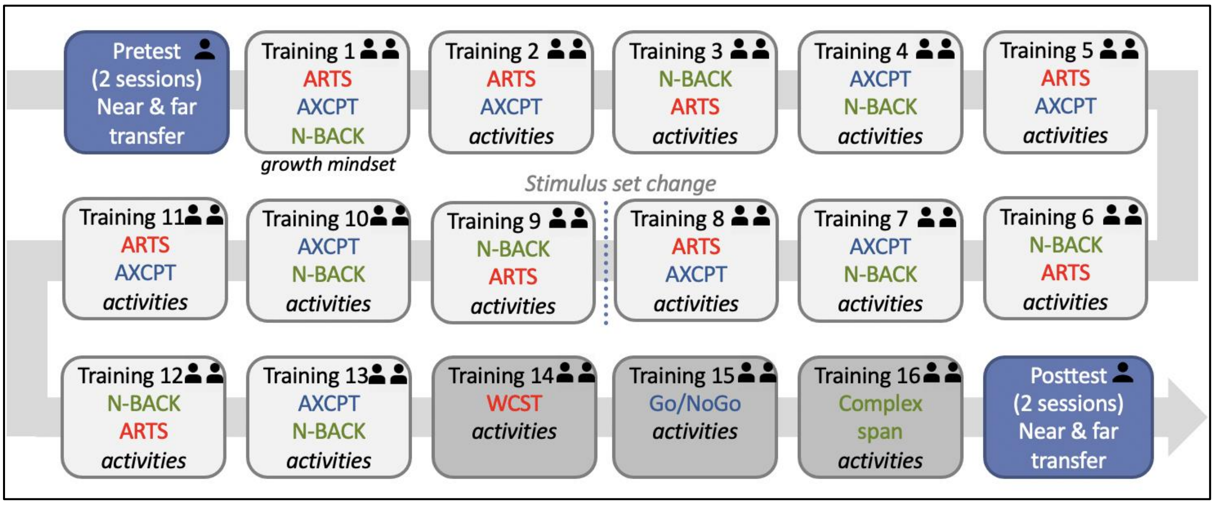


Figure 1. Overall procedure of the study. Pre- and post-test sessions were completed individually and training sessions in pairs. Red, blue, and green indicate set-shifting, inhibition, and working memory tasks, respectively. 'Activities' refer to metacognitive or creative exercises. ARTS = Alternating Runs Task Switching, AXCPT = AX-Continuous Performance Task, WCST = Wisconsin Card Sorting Test.

As indicated in Table 2, Meta-EF and Basic-EF groups both took part in adaptive EF exercises, while the Control group participated in matched non-adaptive EF exercises. In terms of activities, the Meta-EF group took part in metacognition exercises, while the Basic-EF and Control groups participated in matched creative exercises.

Table 2. Overview of targeted skills, task difficulty, and activities for the three groups

| | Control | Basic-EF | Meta-EF |
|---|---|---|---|
| Targeted Skills | - | EF | EF + Metacognition |
| Task Difficulty | Non-adaptive | Adaptive | Adaptive |
| Activity | Creative exercise | Creative exercise | Metacognitive exercise |

### 2.4 Training tasks

In this section, we will focus on the EF tasks used for training throughout sessions. Please refer to Figure 2 for the illustrations of the three main EF training tasks (AX-CPT, ARTS, N-back). All the following EF tasks were adaptive for the Meta-EF and Basic-EF groups but not for the Control group. For detailed descriptions (including information on task adaptivity), please refer to the Supplementary Materials.

*AX-Continuous Performance Task* (*AX-CPT).* This task was used to train participants' inhibitory skills throughout training sessions 1-13. The participants were presented with sequences of pictures, including pairs of cues (wand or hat) and probes (cauldron or broom). At probe onset, the child was instructed to press one of two response keys, associated to either target or non-target responses. If the cauldron appeared after the wand, then the participants were instructed to press the green button (key "A"), but if any other order of stimuli occurred (e.g., the wand followed by the broom) then they were told to press the yellow button (key "L"). The task consisted of a practice phase with two blocks (18 trials each) and an experimental phase with two blocks (30 trials each). In terms of adaptivity, for the Meta-EF and Basic-EF groups, the participants all started with level 0 and the level increased or decreased after the participants completed each block. As the levels increased, the time to respond got shorter and the proportion of AX trials increased, making the task more challenging. The level (a) increased if the error rate in that block was below 20% (b) stayed at the same level if the error rate was 20%- 40% (c) decreased if the error rate was below 40%. The error rate was calculated based on all trials of the previous block of non-AX trials. If participants failed to respond in time, this counted as an error as well. For the Control group, block and trial procedures were identical to the other two groups but the task difficulty was not adapted. The blocks always included 50% of AX trials and the participants had 3500-ms to respond.

*Alternating Runs Task Switching* (*ARTS).* This task was used to train set shifting and switching skills throughout training sessions 1-13. Participants performed two tasks (A and B) either in single-task blocks (task A or B separately; Control group) or in mixed-task blocks (switching between both tasks; Meta-EF and Basic-EF groups). In mixed-task blocks, participants were instructed to switch tasks on every second trial. Task A requires participants to decide whether a picture shows a dragon or a troll and task B whether the picture is shown upright or upside down. There were three blocks of practice trials (10 trials each) and four blocks of experimental trials (17 trials each). For the Meta-EF and Basic-EF groups, the stimulus presentation and the response time window decreased as performance improved and vice versa. In similar vein to the AX-CPT task, the level (a) increased if the error rate in that block was below 20% (b) stayed at the same level if the error rate was 20%- 40% (c) decreased if the error rate was below 40%. The Control group did the single task blocks only and the stimulus presentation time was always set to 1500-ms and they had a time window of 5000-ms to respond.

*Visuospatial N-back Task (N-back).* This task was used to train the participants' working memory and updating skills throughout training sessions 1-13. Stimuli were presented in a 3x3 grid and the center grid was only used for task prompts (i.e., fixation cross). The stimulus moved from one place to another on the 3x3 grid and the participants were asked to press the space bar every time they see the stimulus in the same place as *n* trials back. There were two practice blocks (one 1-back and one 2-back with 15 trials each) and three experimental blocks (20 + *n* trials each). The Control group always stayed at 2-back level throughout the three experimental blocks, while the Meta-EF and Basic-EF groups adapted the level depending on their performance. Their performance was rated based on the Pr score, which is calculated by subtracting number of False Alarm from Hit trials. The level

(a) increased if the Pr score was more than 4 (b) stayed the same if the Pr score was 3 or 2 (c) decreased if the Pr score was less than 2.

*Wisconsin Card Sorting Test.* This task was used to train set shifting and switching skills in training session 14. Four key cards, numbered 1 to 4, were presented at the top of the screen and response cards were presented one at a time below. The task required participants to match the series of response cards with any of four key cards by pressing the number corresponding to that key card. Response cards were matched by colour (red, green, blue, yellow), shape (triangle, star, cross, circle), or number (1, 2, 3, 4). There were one practice block (20 trials) and two experimental blocks (64 trials each). In the beginning, the sorting rule is switched after 10 consecutive correct sorts, but the frequency of rule switches increased and the time to response decreased as the difficulty level increased for the two training groups. The levels increased if the category was completed with less than 4 errors, stayed the same if it was completed with 4-7 errors, and levelled down if it was completed after more than 8 errors. Each block started with the last difficulty level reached in the previous block. In the control version, the task was exactly the same but the rule changes were explicitly announced and the sorting rule was displayed on the screen simultaneously with the stimulus.

*Go-No-Go task*. This task was used to train inhibition skills in training session 15. In this task, pictures appeared consecutively in the middle of the screen. There were four sets of "go" stimuli and one set of "no-go" stimulus. The children were instructed to press the response key every time a "go" stimulus (i.e., ordinary books, 80% of trials) was presented and withhold from responding every time a "no-go" (i.e., magic book; 20% of trials) stimulus was presented. There was one practice block (10 trials, two "no-go" trials) and five experimental blocks (30 trials each, six "no-go" trials). The level went up or down within blocks; the level increased after 2 consecutive correct responses on "no-go" trials (i.e., no response) and decreased after two consecutive errors on “no-go” trials; otherwise, the level stayed the same. Each block starts with the last difficulty level reached in the previous block. The control version included only “go” stimulus (absence of “no-go” trials).

*Complex spatial span task.* This task was used to train WM skills in training session 16. In this task, the children performed a visuospatial WM task against a secondary processing task. In the encoding phase, WM-stimuli were presented in a 3x3 grid (one at a time, starting with two items to be remembered, i.e., span = 2). After each stimulus, the children performed the processing task (react to a stimulus by clicking on one of two response buttons presented under the grid). In the recall phase, participants were required to recall the position of the WM stimuli in the correct order. There was one practice block (10 trials with span = 2) and three experimental blocks (20 trials each). The task started with 2 items to be remembered (span = 2) and the span increased as performance improved and vice versa for the two training groups. The level went up after 2 consecutive correct responses; went down after two consecutive errors; and otherwise stayed at same level. In the control version, the task difficulty did not change, and the task difficulty was always set to span length of 2. Each block started with the last difficulty level reached in the previous block.

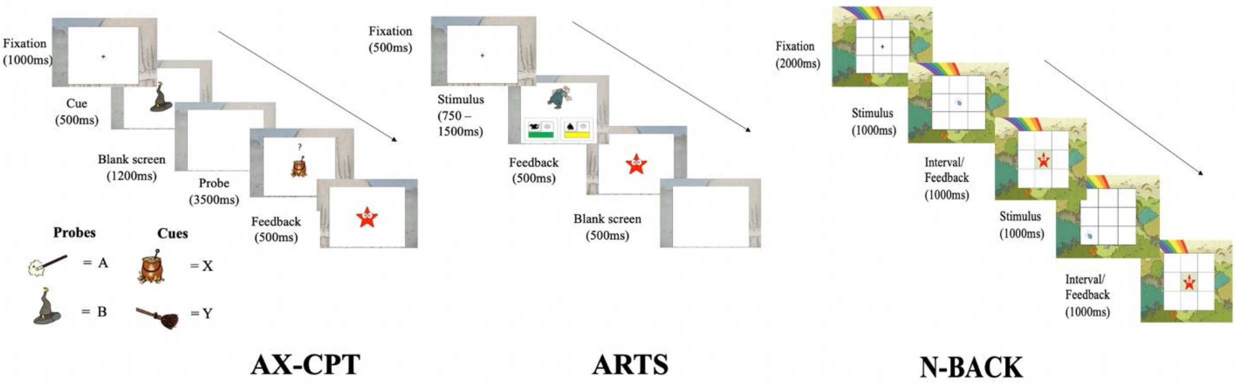


Figure 2. Illustrations of AX-CPT (left), ARTS (middle), and N-back (right). AX-CPT = AX-Continuous Performance Task, ARTS = Alternating Runs Task Switching.

**2.5 Near Transfer Tasks**

In this section, we describe EF tasks that are used to examine near-transfer effects of trained components. The illustrations of these tasks are presented in Figure 3.

*Antisaccade task*. In order to locate the target briefly presented on the other side of the monitor, participants must resist the reflexive urge to look at a visual cue that appears unexpectedly on the peripheral field of the screen and instead look in the opposite direction. First, a fixation cross was shown in the center of the panel followed by a cue (i.e., goblin) on one side of the screen, then a target on the other side. The targets were selected at random from a list of target stimuli (dog, pig, donkey, goose), with the restriction that they appeared equally frequently. The mask (i.e., tree) appeared to cover the target stimulus after 100-ms of target presentation, and the participants were given a maximum of 1500-ms to respond. The task started with an instruction, then two blocks of practice trials (eight trials each), and finally three blocks of experimental trials (24 trials for each block). The child was seated precisely 60 cm away from the monitor, as sitting further away from the monitor would make identifying the animal easier even after fixating on the cue. The accuracy of the experimental trials was used for further regression analyses.

*Cued task switching (CTS).* Participants completed two tasks (A and B) in single-task blocks (task A or B separately) or mixed-task blocks (switching between task A and B). Four practice blocks (two single blocks x five trials and two mixed blocks x 10 trials) and four single blocks and six mixed blocks of experimental trials were used (17 trials each). Participants switched tasks in mixed-task blocks based on a visual cue that indicated shape- (i.e., an image of a flower or a tree) or size-sorting (i.e., an image of a small circle or a big circle). In the mixed block, half of the trials were switch trials (i.e., the task shifted from the previous trial) and the other half were stay trials (i.e., the task remained the same). Additionally, in half of the trials (for both single and mixed blocks), the cue was presented 1000-ms before the stimulus appeared, making proactive preparation ahead of stimulus onset possible (i.e., cue is visible). In the other half, the cue was masked and presented at the same time as the stimulus. "XX" appeared on the screen 1000-ms before the presentation of the stimulus, rather than a cue, rendering proactive preparation impossible (i.e., cue is not visible). The difference in performance between stay trials and switch trials in mixed-task blocks (i.e., switch costs) as well as the difference in performance between stay trials in mixed-task blocks and single trials in the single blocks (i.e., mixing costs) were used for further regression analyses as proxies for cognitive flexibility. Furthermore, proactive control was measured as the difference in reaction times (RTs) between "cue visible" trials and "cue not visible" trials in the mixed blocks. Higher values indicate better proactive control.

*Backwards Corsi-block task.* The children were asked to remember the sequence of moves in reverse order after a stimulus moved in a 3x3 grid on the computer screen. A stimulus (i.e., cat) was introduced one by one at a random location in the 3x3 grid in each trial. The task started with two grid locations. For each sequence length, six trials were presented, with each sequence length increasing by one location if the participant got at least two trials correct. The maximum sequence length that the participants could reach was 8. One block of demo trials and one block of practice trials (10 trials each) were included in this task, with a maximum of seven blocks of experimental trials (six trials each). The product score, which is the product of the sequence span and the number of correctly remembered trials, was used to assess working memory capacity. This composite calculation considers both the correct number and the span length, making it a more accurate and fine-grained metric than the maximum span length reached (Kessels et al., 2000).

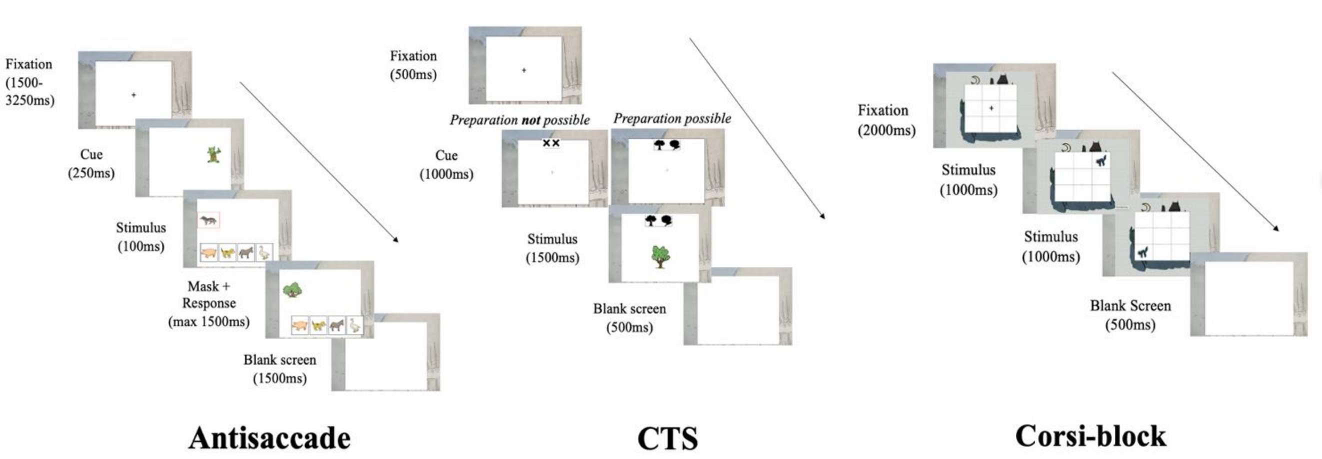


Figure 3. Illustrations of Antisaccade (left), CTS = Cued Task Switching (middle), and Corsi-block (right) tasks.

**2.6 Far Transfer Tasks**

We now describe far-transfer tasks that assessed the effects of training to untrained components beyond EF, namely, academic skills (i.e., intelligence, reading, and math).

*Matrix Reasoning task*. We used a subtest of the Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2004). This task assessed non-verbal reasoning. This was a nonverbal task in which one had to complete geometrical figures by selecting the correct missing piece from five options. After three successive failures, the task was discontinued.

Scores corresponded to the sum of correct responses (1 point for each trial) with a maximum score of 32. Children all started from item one regardless of age.

*Reading comprehension task.* We used a different subset of items from the subtest of the Wechsler Individual Achievement Test Third Edition (WIAT-III; Wechsler, 2017) in each of the pre- and post-test sessions. The children were asked 6-7 content-related questions for each of the two short stories in this task. With a maximum score of 26, scores corresponded to the number of correct responses (13 questions total, one point for partial and two points for complete answers).

*Math problem solving task.* We used a subset of items from the subtest of the Wechsler Individual Achievement Test Third Edition (WIAT-III; Wechsler, 2017) in each session. The children were required to solve a maximum of 25 math questions (e.g., basic concepts, everyday applications, geometry, and algebra). The task was aborted when the children made three consecutive errors. Scores corresponded to the sum of correct responses (one point for each trial) with a maximum score of 25.

#### *2.5.2 Activities*

As indicated in Table 3, the Meta-EF group underwent metacognitive activities along with adaptive EF tasks, while the Basic-EF and Control groups completed activities related to fostering creativity, positive thinking, and self-respect along with adaptive (Basic-EF) and non-adaptive (Control) EF tasks. We created two workbooks, one for the Meta-EF group and one for the Basic-EF and Control groups. The workbooks, specific instructions and scripts for each session, as well as detailed session aims can be accessed through the Supplementary Materials.

The metacognitive activities targeted mostly metacognitive monitoring and, to a lesser extent, metacognitive control, with the aim to foster metacognitive reflection on executive function engagement. They were completed in pairs to encourage children to exchange their respective reflections. Specifically, they targeted task demand identification, strategy reflection and monitoring, and performance evaluation. Task demand identification activities were aimed at promoting the evaluation of specific task demands and sources of interference. For example, participants were asked to identify the sources of interference (i.e., say why each task is difficult) and indicate, among different types of trials (e.g., congruent vs. incongruent) or different versions of a task (e.g., 1-back vs. 2-back), which is easier or more difficult and why. They also watched videos where Leo and Lessa (i.e., the protagonists of the cover story) explain how they use different metacognitive strategies to approach each task and were encouraged to model their behaviour. In other sessions, they had to predict to what extent proposed task modifications (e.g., repeating the stimuli more or less frequently in an N-Back task) would change task difficulty, and to rate the difficulty of each task before and after completing it. Strategy reflection activities were designed to promote reflection on how one completes a task, on the costs and benefits of possible strategies, and generate and select a strategy for upcoming tasks. Participants were instructed to "coach" the experimenter by explaining the task structure and how they approach the task. They also completed exercises in which they consider alternative strategies suggested by the experimenter and highlight the potential costs and benefits of each. The performance evaluation activities were aimed at promoting performance monitoring and evaluating the fit of a specific strategy to the task demands. Participants were invited to generate their own feedback, for instance by choosing between two performance scores (e.g., a clock indicating correct but too late responding, and a thumb up signalling correct responding in time) that accurately reflects their performance.

The Basic-EF and the Control group instead undertook exercises that focused on fostering empathy, group work, growth mindset, and creativity. The activities were organized in a similar manner as the Meta-EF group, in which the participants worked on various tasks indicated in the workbook in pairs for 15 minutes in total (three exercises for 5 min each). For instance, the participants were asked to identify different groups that they belong to and express how they felt about being a member of each group. In other sessions, they were asked to reflect on their feelings and engage in calming activities such as colouring and drawing. Growth mindset was an activity that all groups underwent, where they watched several videos that introduced this concept and discussed certain phrases/languages that can be used to foster one's growth mindset.

Table 3. Summary of the activities in each training group. "Component" refers to the specific component of metacognition/creativity skills that was targeted in each session and "Activities" refer to brief descriptions of the activities in each Component.

| | Component | Activities |
|---|---|---|
| | **Sessions 1-4** | |
| **Meta-EF** | | |
| | Monitoring | • Words to describe uncertainty<br>• Uncertainty judgements using the sure-o-meter (individually and in pairs) |
| | Task Demands | • Ordering task instructions<br>• Describing the task<br>• Sequencing language |
| | Growth Mindset | • Video 1 - introducing a growth mindset<br>• Video 2 - the value of making mistakes<br>• Video 3 - growth mindset language and the incredible power of 'yet' |
| | Monitoring | • Words to describe uncertainty<br>• Uncertainty judgements using the sure-o-meter (individually and in pairs) |
| **Control + Basic-EF** | | |
| | Similarities between individuals | • Sharing likes / dislikes and looking for similarities |
| | Benefits of a group | • Identifying how teams work together<br>• Recognising what groups they're in<br>• Thinking about how being in a group feels |
| | Growth Mindset | • Video 1 - introducing a growth mindset<br>• Video 2 - the value of making mistakes<br>• Video 3 - growth mindset language and the incredible power of 'yet' |
| | **Sessions 5-7** | |
| **Meta-EF** | | |
| | Monitoring task demands | • Identifying when an elf character misses a step in the instructions |

| | | |
|---|---|---|
| | Strategy Use | • Words to describe strategies<br>• Linking different strategies to different games and different people (predicting which 'thinking word' the elves will use for different games)<br>• Identifying the strategies when seen in the videos<br>• Personalising strategies by identifying how we feel when we use each strategy |
| | Goal Maintenance | • Maintaining a goal on a physical post-it note<br>• Maintaining a goal on a mental post-it note |
| | Growth Mindset | • Separating growth mindset phrases from fixed mindset phrases<br>• Challenge to use favourite growth mindset phrase during session |
| **Control + Basic-EF** | | |
| | Empathy | • Watching Empathy videos<br>• Discussing how to be empathic |
| | Recognising strengths | • Identify what you'd like to work on<br>• Finding a skill you're proud of |
| | Appreciating imperfection | • Watching 'ish' story<br>• Describing 'ish' pictures<br>• Drawing favourite things about self 'ish' |
| | Growth Mindset | • Separating growth mindset phrases from fixed mindset phrases<br>Challenge to use favourite growth mindset phrase during session |
| | **Sessions 8-10** | |
| **Meta-EF** | | |
| | Goal calibration | • Choosing a goal for each block<br>• Goal maintenance |
| | Strategy use | • Linking strategies to 'new' games (transfer)<br>• Choosing which strategies will help to achieve goals |
| | Task demands | • Identifying difficulties in games (from elves and self)<br>• Choosing strategies to solve these problems |
| | Generalising | • Linking problems in games to real life and discussing if the same strategies will help |
| **Control + Basic-EF** | | |
| | Reflecting on feelings | • Thinking about how colours make us feel<br>• Drawing with colours that make you feel good |
| | Calming activities | • Colouring<br>• Reflecting on activities that make you feel calm<br>• Creating mandalas |
| | **Sessions 11-13** | |
| **Meta-EF** | | |
| | Planning | • Planning vocabulary<br>• Making own plans using GOT IT planning, deciding goals, obstacles, thinking words |
| | Monitoring | • 'Stop' step of GOT IT planning<br>• 'Is it working' step of GOT IT planning |

| | | |
|---|---|---|
| | | • Observe elves' plans and predict who will do better |
| | Evaluating with growth mindset | • Talk nicely<br>• Tickboxes after EF game to assess other gains |
| | Generalising | • Planning for a real-world task<br>• Planning for the new games that will be played |
| **Control + Basic-EF** | | |
| | Importance of perspective | • Draw half happy half sad, tear and throw away sad<br>• Colouring with positive colours<br>• Make two faceless sticks "happy" and "sad" with tiny change<br>• Happiness shield |
| | Focusing on the positive | • The dot<br>• Together collage sad colours, then cover with happy |
| | Reframing negative feelings | • Find positive things to say about sad situations<br>• Make scary pictures funny |
| | **Sessions 14-16** | |
| **Meta-EF** | | |
| | Planning | • Making Stop GOT IT plans after first round of game |
| | Monitoring | • Refer to plan |
| | Task Demands | • Explain games using sentence starters |
| | Problem-solving | • Articulate questions about the games and try to answer them in the next round |
| | Generalisation | • Interview partner about games |
| **Control + Basic-EF** | | |
| | Expressing yourself appropriately | • Recognising inappropriate monster and mouse<br>• How to ask for what you want assertively<br>• Recognising when you act inappropriately<br>• Thinking about how inappropriate expression affects others |
| | Focusing on strength | • Drawing your strengths |
| | Nurturing growth | • Sharing good ideas you've had<br>• Thinking about how to help ideas grow |

### 2.6 Data Analysis

We ran linear mixed effects models and generalized linear mixed effects models for all analyses below (i.e., training gains, executive function gains, academic performance gains) by using the lme4 package (Bates et al., 2015) in R (version 4.1.0). Before running the models, reaction time data were all log-transformed. In all models, we included Session (*pre*, post), Group (*Meta-EF*, Basic-EF, Control) and its interactions as fixed effects, Country/Testing site (*Germany*, UK) as a covariate and Subject as a random intercept. For the analysis on proactive control index and cognitive flexibility in the Cued Task Switching, we also included Cue_Visibility (*no: proactive control **not** possible*, yes: proactive control possible) and Trial_Type (*stay*, single, switch) respectively as condition variables and their three-way interaction with Session and Group. These categorical variables were treatment coded and the levels indicated in italic were set as the reference level. *P*-values were obtained by likelihood ratio tests of the null model against the model with the effect in question. We ran a pairwise Tukey's test comparison using the emmeans package in R (Lenth, 2021) to examine whether (a) there were baseline differences among the groups and (b) whether there were group differences in performance gains from pre- to post-test.

## 3. Results

### 3.1 Training gains

Since the difficulty level increased as participants performed better on the training tasks, reaction time and error measures are not appropriate to assess gains across training sessions. Thus, we compared the difficulty level that the participants reached at the beginning and end of training sessions for each task. We will only report training gains from the Basic-EF and Meta-EF groups, since the difficulty level for the Control group always stayed the same across sessions. Figure 4 visualizes the training gains for the three tasks.

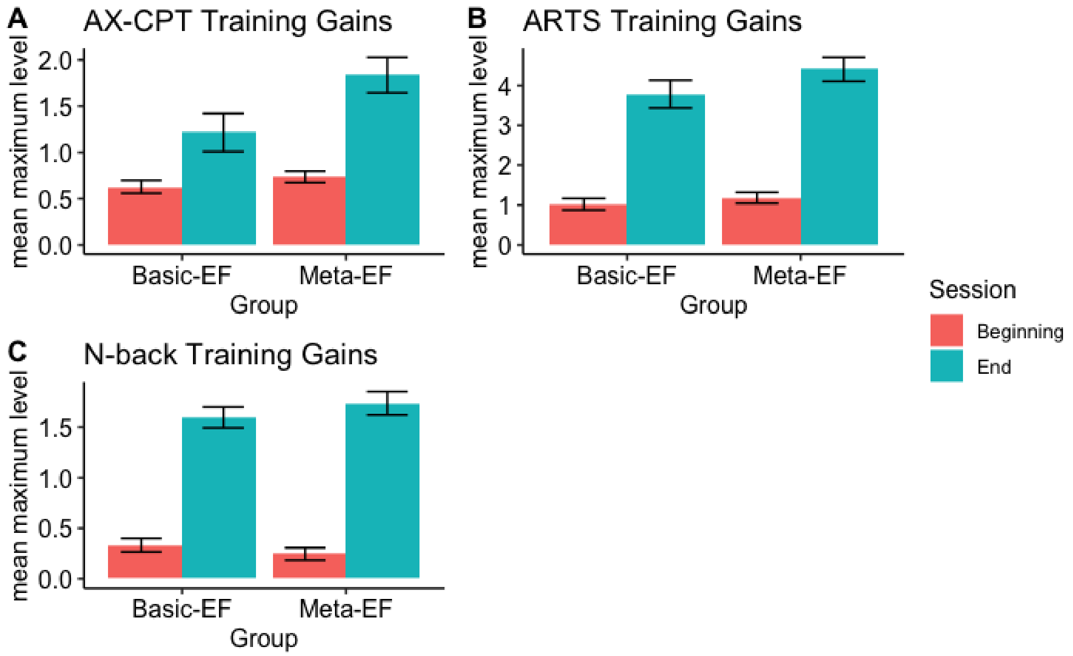


Figure 4. Training gains of the three tasks split by Session (beginning, end) and Group (BasicEF, MetaEF). Panel A: Training gains in AX-CPT; Panel B: Training gains in ARTS; Panel C: Training gains in N-back. Error bars indicate standard errors. AX-CPT = AX-Continuous Performance Task, ARTS = Alternating Runs Task Switching.

### *3.2.1 AX-CPT*

As shown in Figure 4 Panel A, there was a significant main effect for Session ($E = .57$, $p = .002$) but no significant effect for Group ($E = .10$, $p = .60$) and a near-significant interaction between Session and Group ($E = .51$, $p = .05$). This means that both groups increased their difficulty level to a significant degree and the Meta-EF group tended to improve more than the Basic-EF group. There was no significant effect of Country ($E = -.02$, $p = .87$), which indicated no differences in performance between those tested in UK and in Germany.

### *3.2.2 ARTS*

There was a significant main effect for Session ($E = 2.75$, $p < .001$) but no significant effects for Group ($E = .16$, $p = .63$) nor interaction between Session and Group ($E = .45$, $p$

= .31). Thus, difficulty level increased over time at a similar rate in both groups in the ARTS task (Figure 4, Panel B). There was no significant effect of Country ($E$ = .21, $p$ = .42).

#### *3.3.3 N-back*

The output showed a significant main effect for Session ($E$ = 1.25, $p$ < .001) but no significant effects for Group ($E$ =-.08, $p$ = .49) or interaction between Session and Group ($E$ = .22, $p$ = .16). Children, regardless of whether they were in the Meta-EF or the Basic-EF group, improved their level to a significant extent over time (Figure 4, Panel C). There was no significant effect of Country ($E$ = -.14, $p$ = .15).

### 3.2 Baseline differences at pre-test

There were significant differences in proactive control index between the Basic-EF and Control group, with the Controls showing greater proactive control engagement than the Basic-EF group at pre-test ($E$ = -.31, $p$ = .004). Otherwise, there were no significant differences between groups on all measures of executive function and academic skills at pre-test.

### 3.3 Near transfer effects

Here, we report the near transfer effects on untrained mesures of EF processes: Cued Task Switching (CTS), Backwards Corsi-block, and Antisaccade from pre- to post-test.

#### *3.3.1. CTS*

Table 4 shows the model using the Type III Analysis of Variance Table with Satterthwaite's method output for reaction time and accuracy on the Cued Task Switching task. There were significant main effects of Group, Trial Type, Session, as well as an interaction between Group and Session for reaction time data. As for accuracy, there were significant main effects of Group, Trial Type, Session, as well as interactions between Group and Session as well as Trial type and Session. No three-way interactions between Group, Trial Type, and Session were found for either reaction time (RT) or accuracy models. Thus, changes in switch and mixing costs for both RT and accuracy did not differ across the three groups (see Figure 5). Moreover, simple pairwise contrasts between Session and Group for the RT model revealed that the Meta-EF group showed a greater RT decrease from pre-test to post-test than both Basic-EF ($E$ = .08, $p$ < .001) and Control groups ($E$ = .29, $p$ < .001). Moreover, the Basic-EF group became faster at responding from pre- to post-test than the Control group ($E$ = .20, $p$ < .001). In contrast, accuracy decreased from pre-test to post-test in the Meta-EF ($E$ = .33., $p$ < .001) and Basic-EF groups ($E$ = .30, $p$ < .001), with no difference between the two groups ($E$ = .03, $p$ = .61), whereas it stayed consistent in the Control group ($E$ = -.001, $p$ = .97). Simple pairwise contrasts between Trial type and Session further revealed an overall reduction in mixing costs ($E$ = -.39, $p$ < .001) from pre- to post-test but not in switch costs ($E$ = .009, $p$ = .87) for accuracy.

Table 4. Linear mixed effects regression model output for switch costs and mixing costs (Reaction Time and Accuracy) on Cued Task Switching

| **RT** | **Sum Sq** | **Mean Sq** | **F value** | **P-value** |
|---|---|---|---|---|
| Group | 3.39 | 1.70 | 12.20 | **< 0.001** |
| Trial_type | 228.73 | 114.36 | 822.83 | **< 0.001** |
| Session | 267.40 | 267.40 | 1923.92 | **< 0.001** |
| Age | 1.68 | 1.68 | 12.11 | **< 0.001** |
| Country | 0.06 | 0.06 | 0.46 | 0.50 |
| Group:Trial_type | 0.91 | 0.23 | 1.63 | 0.16 |
| Group:Session | 132.43 | 66.21 | 476.40 | **< 0.001** |
| Trial_type:Session | 0.82 | 0.41 | 2.97 | 0.05 |
| Group:Trial_type:Session | 2.42 | 0.61 | 4.35 | < 0.001 |

| **Accuracy** | **Sum Sq** | **Mean Sq** | **F value** | **P-value** |
|---|---|---|---|---|
| Group | 13.01 | 6.51 | 6.51 | **0.003** |
| Trial_type | 408.77 | 204.38 | 204.38 | **<.001** |
| Session | 75.34 | 75.34 | 75.34 | **<.001** |
| Age | 5.07 | 5.07 | 5.07 | **.03** |
| Country | 0.05 | 0.05 | 0.05 | 0.43 |
| Group: Trial_type | 5.57 | 1.39 | 1.39 | 0.13 |
| Group:Session | 31.83 | 15.91 | 15.91 | **<.001** |
| Trial_type:Session | 58.60 | 29.30 | 29.30 | **<.001** |
| Group: Trial_type:Session | 1.65 | 0.41 | 0.41 | .79 |

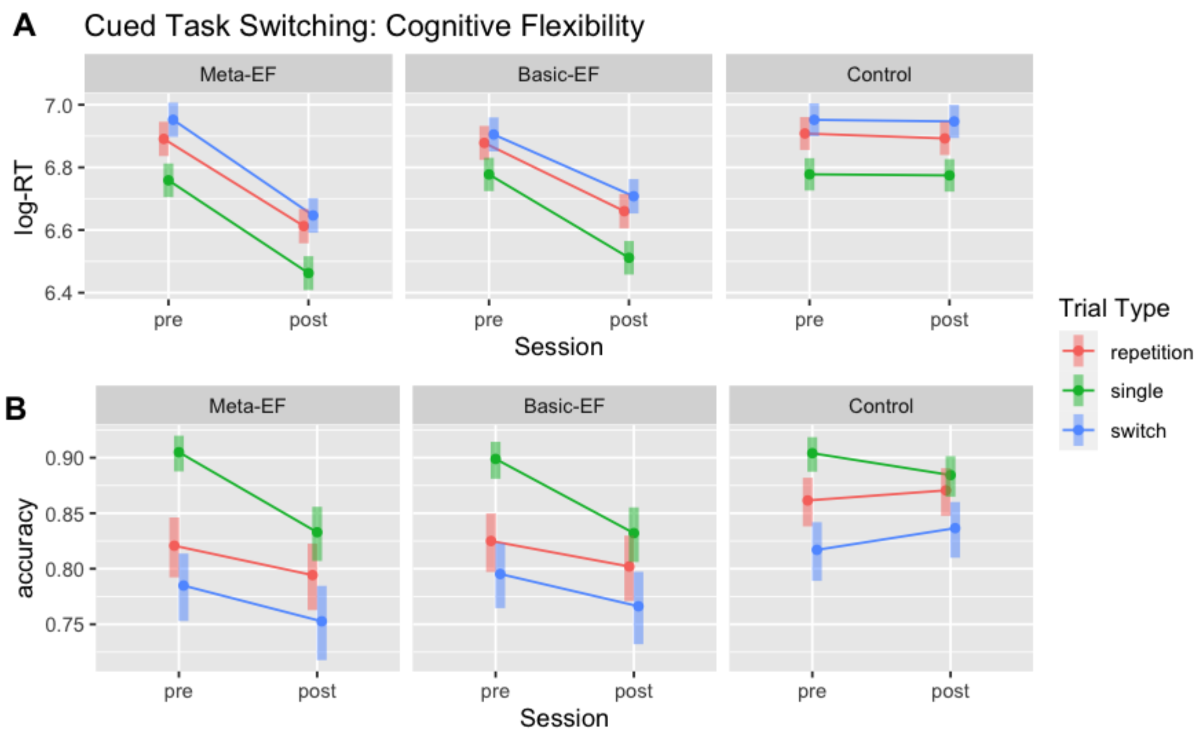


Figure 5. Predicted performance gains on the Cued Task Switching task split by Trial Type (repetition, single, switch), Group (Basic-EF, Meta-EF, Control), and Session (pre, post) for reaction time (top A panel) and accuracy (bottom B panel) data. Error bars indicate confidence interval.

Regarding the proactive control scores (i.e., proactive not possible RT – proactive possible RT), as shown in Table 5, there were significant main effects of Group, Session, Cue Visibility, a significant interaction between Group and Session, Session and Cue Visibility, and a three-way interaction between Group, Session, and Cue_visibility. Pairwise contrasts of the three-way interaction indicates that both Meta-EF ($E$ = -.05, $p$ = .02) and Basic-EF groups ($E$ = -.08, $p$ < .001) engaged control more proactively over time than the Control group (see Figure 6 for visualization), with no difference between the Meta-EF and Basic-EF groups ($E$ = -.03, $p$ = .23). Results from the pairwise comparison between Group and Session indicate that Meta-EF group's global RT decreased to a greater extent than the Basic-EF ($E$ = .08, $p$ < .001) and Control ($E$ = .28, $p$ < .001) and Basic-EF became faster at responding from pre- to post-test than the Control group ($E$ = .20, $p$ < .001).

Table 5. Linear mixed effects regression model output for proactive control index on Cued Task Switching

| | **Sum Sq** | **Mean Sq** | **F value** | **P-value** |
|---|---|---|---|---|
| Group | 2.64 | 1.32 | 9.73 | < **0.001** |
| Session | 138.22 | 138.22 | 1017.22 | < **0.001** |
| Cue_visibility | 220.95 | 220.95 | 1626.11 | < **0.001** |
| Age | 1.30 | 1.30 | 9.59 | < **0.001** |
| Country | 0.11 | 0.11 | 0.84 | 0.36 |
| Group:Session | 74.74 | 37.37 | 275.01 | < **0.001** |
| Group: Cue_visibility | 0.15 | 0.07 | 0.54 | 0.59 |
| Session: Cue_visibility | 4.41 | 4.41 | 32.42 | < **0.001** |
| Group:Session: Cue_visibility | 1.68 | 0.84 | 6.20 | < **0.001** |

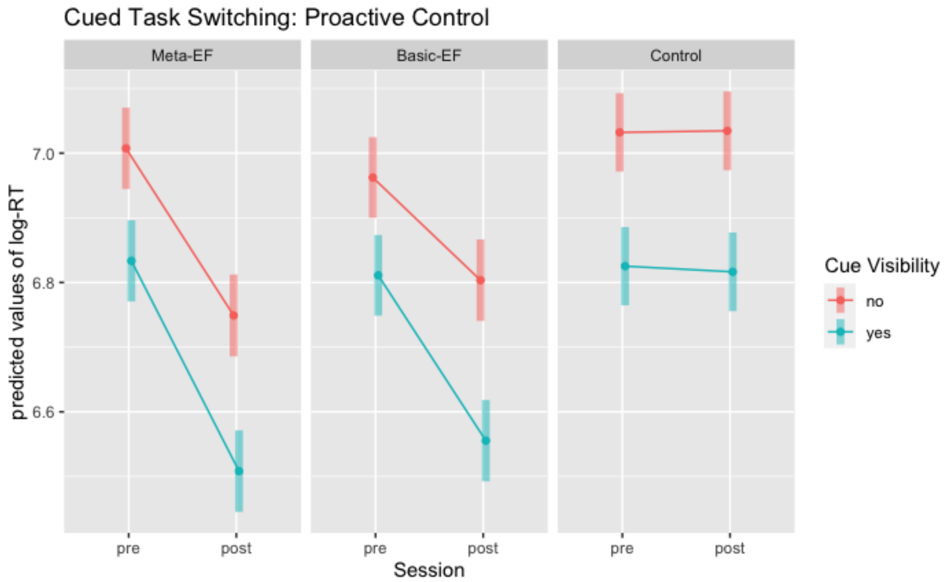


Figure 6. Predicted performance gains on the Cued Task Switching task split by Cue_visibility (yes = cue visible, no = cue **not** visible), Group (Basic-EF, Meta-EF, Control), and Session (pre, post) for reaction time data. Error bars indicate confidence interval.

### *3.3.2. Backwards Corsi-block*

As for the working memory product score on the Backwards Corsi-block task, there was a significant effect of Session as well as a Group and Session interaction (see Table 6). Working memory product scores improved to a greater extent in both the Meta-EF (*E* = -12.41, *p* = .003) and Basic-EF groups (*E* = -8.93, *p* = 0.03) than the Control group (see Figure

7 for visualization). No significant difference in working memory product score gains was found between the Meta-EF and Basic-EF groups ($E$ = -3.49, $p$ = 0.39).

Table 6. Linear mixed effects regression model output for working memory product score on Backwards-Corsi-block task

| | **Sum Sq** | **Mean Sq** | **F value** | **P-value** |
|---|---|---|---|---|
| Group | 20.48 | 10.24 | 0.06 | 0.94 |
| Session | 4374.69 | 4374.69 | 26.13 | < **0.001** |
| Age | 781.51 | 781.51 | 4.67 | **0.03** |
| Country | 783.23 | 783.23 | 4.68 | **0.03** |
| Group:Session | 1634.75 | 817.38 | 4.88 | **0.01** |

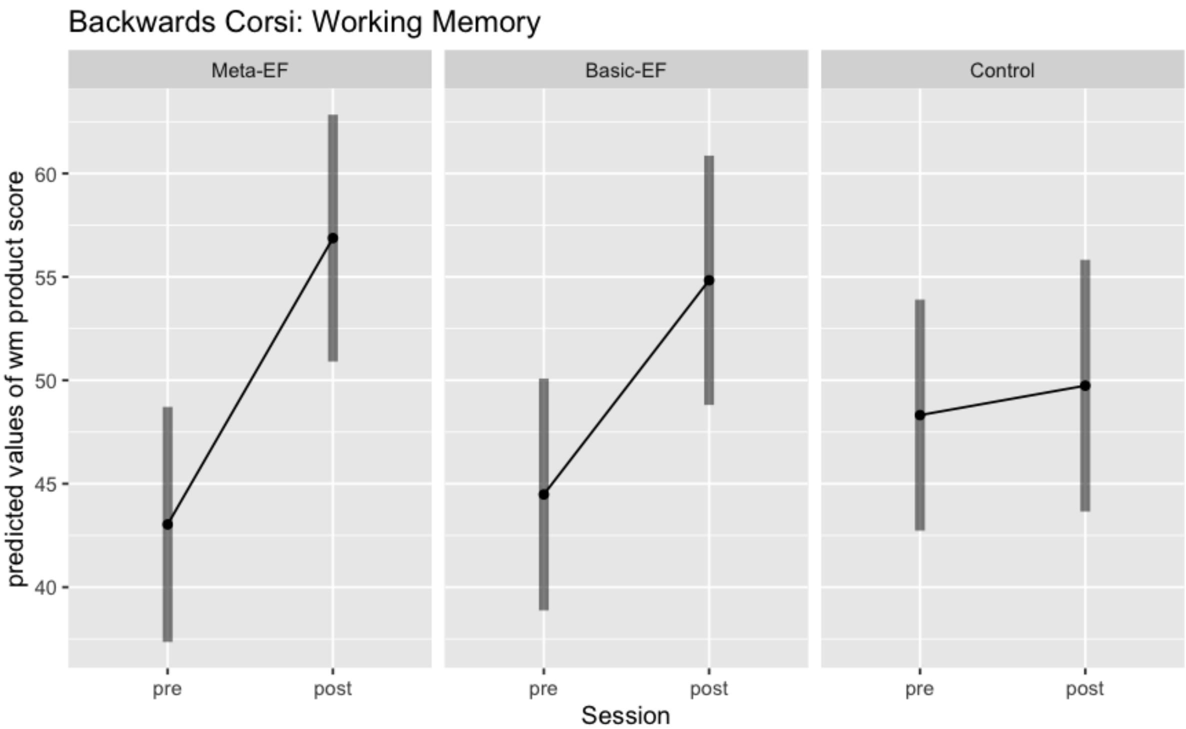


Figure 7. Predicted performance gains on the Backwards Corsi-block task split by Group (Basic-EF, Meta-EF, Control) and Session (pre, post) for working memory product score. Error bars indicate confidence interval.

### *3.3.1. Antisaccade*

For the Antisaccade task, there was a significant effect of Session and a near-significant interaction between Group and Session (see Table 7). The pairwise contrasts revealed significant difference in accuracy gains from pre to post-test between Meta-EF and Basic-EF ($E$ = -.17, $p$ = 0.02), with Basic-EF improving to a greater extent than Meta-EF. No differences in improvements in inhibition were found between Meta-EF and Control ($E$ = .06, $p$ = .39) or Basic-EF and Control ($E$ = -.11, $p$ = .14) (see Figure 8 for visualization).

Table 7. Generalized linear mixed effects regression model output for accuracy on Antisaccade task

| | **Sum Sq** | **Mean Sq** | **F value** | **P-value** |
|---|---|---|---|---|
| Group | 3.08 | 1.54 | 1.54 | 0.22 |
| Session | 173.13 | 173.13 | 173.13 | < **0.001** |
| Age | 9.88 | 9.88 | 9.88 | **0.001** |
| Country | 1.23 | 1.23 | 1.23 | 0.80 |
| Group:Session | 5.39 | 2.69 | 2.69 | 0.06 |

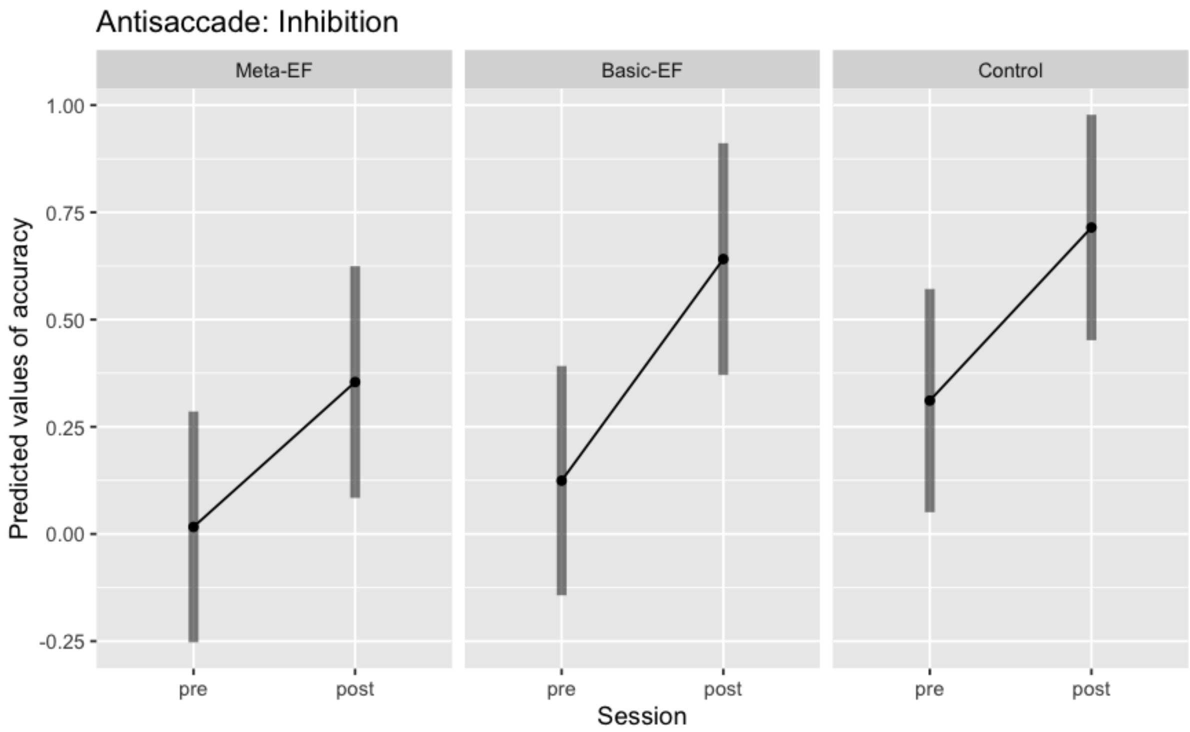


Figure 8. Predicted performance gains on the Antisaccade task split by Group (Basic-EF, Meta-EF, Control) and Session (pre, post) for accuracy. Error bars indicate confidence interval.

### 3.4 Far transfer effects

We will now report the results of the far transfer effects, namely, the development of math, reading, and matrix reasoning scores from pre- to post-test. The output of these models, as indicated in Table 8, informs us about whether training effects of the three academic skills differ among the three groups.

Table 8. Output of the linear regression model for math, reading, and matrix reasoning scores.

| Predictors | **Sum Sq** | **Mean Sq** | **F value** | **P-value** |
|---|---|---|---|---|
| **Math** | | | | |
| Group | 18.87 | 9.43 | 0.88 | 0.41 |
| Session | 181.66 | 181.66 | 17.02 | < **0.001** |
| Country | 9.39 | 9.39 | 0.88 | 0.35 |
| Age | 242.53 | 242.53 | 22.72 | < **0.001** |
| Group:Session | 4.92 | 2.46 | 0.23 | 0.79 |
| **Reading** | | | | |
| Group | 183.44 | 91.72 | 2.84 | 0.06 |
| Session | 135.55 | 135.55 | 4.20 | **0.04** |
| Country | 131.65 | 131.65 | 4.08 | **0.04** |
| Age | 694.60 | 694.60 | 21.52 | < **0.001** |
| Group:Session | 50.61. | 25.31 | 0.78 | 0.46 |
| **Matrix Reasoning** | | | | |
| Group | 19.94 | 9.97 | 0.78 | 0.46 |
| Session | 178.98 | 178.98 | 14.06 | < **0.001** |
| Country | 56.24 | 56.24 | 4.42 | **0.04** |
| Age | 79.66 | 79.66 | 6.26 | **0.01** |
| Group:Session | 23.80 | 11.90 | 0.93 | 0.39 |

Table 8 shows that Session was the only significant variable for all three academic skills. No interactions between Session and Group were found for any of the academic skills (see Figure 9 for visualization). There was a significant effect of Country on matrix reasoning performance, whereby the children from Germany performed better than the ones from the UK. Although the interactions were not significant, we examined the improvement in each group by running post-hoc pairwise contrasts, as these were part of our a priori hypothesis. The pairwise contrasts revealed that Meta-EF group showed significant performance increase in math ($E$ = -2.13, $p$ = .003), reading ($E$ = -2.84, $p$ = .03), and matrix reasoning ($E$ = -2.43, $p$ = .0013) scores, while the Control and Basic-EF groups showed significantly improvement on math only (Control: $E$ = -1.56, $p$ = .02; Basic-EF: $E$ = -1.36, $p$ = .06). The effect size of the intervention effect for the Meta-EF group calculated by using the eff_size function in emmeans (Lenth, 2021) was the largest for matrix reasoning ($g$ = .64), followed by math ($g$ = .58), and reading ($g$ = .50) (effect size of intervention effect for math was .39 for Control and .36 for Basic-EF).

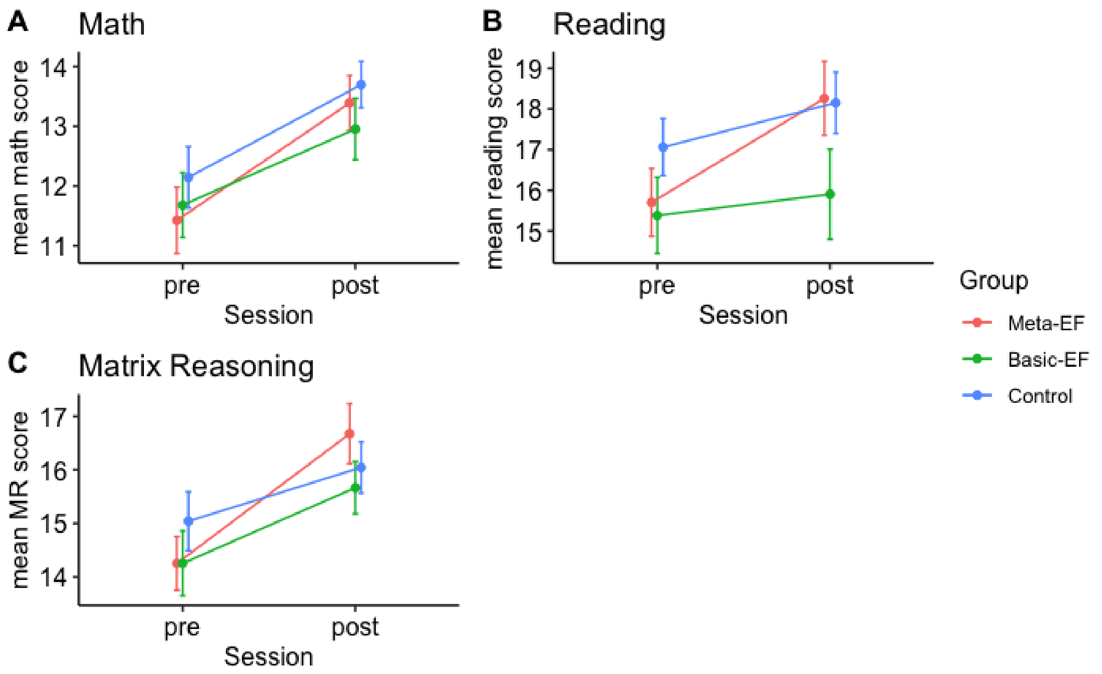


Figure 9. Plot graphs of math (Panel A), reading (Panel B), and matrix reasoning (Panel C) performance split by Session and Group. Error bars indicate confidence interval.

## 4. Discussion

We examined whether EF training combined with metacognitive skills training would enhance EF and academic performance, including proactive control engagement, in children from disadvantaged backgrounds, to a greater extent than EF training alone or no adaptive training. EF training improved working memory and was associated with greater proactive control engagement, relative to the active control group. Training-elicited change in the way children engaged cognitive control was further evidenced by faster, albeit less accurate, responses in the cued task-switching paradigm, a pattern that was more pronounced when children additionally received metacognitive training. Although Group and Session did not interact for academic skills, we also found significant gains in all three academic skills for the Meta-EF group, whereas gains for the other groups were restricted to maths. We discuss each of the main findings in turn below.

In terms of near-transfer effects (i.e., to untrained EF tasks), our results are broadly in line with previous studies that found a positive effect of EF intervention on targeted EF processes. The two training groups improved significantly more than the Control group on working memory. This finding is consistent with a meta-analysis suggesting that EF training has largest near-transfer effect on working memory (Kassai et al., 2019). Indeed, EF training benefits working memory in typically developing younger children (Bergman Nutley et al., 2011; Blakey & Carroll, 2015; Röthlisberger et al., 2012; Traverso et al., 2015), older children with lower EF (Dunning et al., 2013; Holmes et al., 2009; Re et al., 2015; Wong et al., 2014), and children diagnosed with ADHD (Bigorra et al., 2016; Chacko et al., 2014; Hovik et al., 2013; Klingberg et al., 2005; Kray et al., 2012). Our study extends prior findings to a population of low SES children, which is important given that these children often have EF difficulties and thus are a priority for targeted interventions.

Importantly, EF training promoted proactive control engagement, suggesting a shift to a more mature way of engaging cognitive control. Both the Meta-EF and Basic-EF groups engaged greater proactive control from pre- to post-test when compared to the Control group, in CTS when the cue was provided ahead of the target. To our knowledge, this is the first evidence that proactive control can benefit from EF training in children. This finding is important given the specific contribution of proactive control to academic skills (Kubota et al., 2020; Skau et al., 2022; Wang et al., 2021), and especially remarkable given that our EF training programs did not target proactive/reactive control strategies per se. We suspected that combining metacognitive reflection with EF training may help children better monitor how and how well they engage control (e.g., whether they engage it reactively or proactively, whether they fully or partially prepared for the upcoming task) and how well they performed, hence leading them to better determine whether proactive control is viable for current task demands. This prediction was rooted in previous findings suggesting that metacognitive reflection training elicits more adult-like neural markers of conflict monitoring in 5-year-olds (Pozuelos et al., 2019), and 6-year-olds engage proactive control more when they are encouraged to actively monitor their performance (Hadley et al., 2020). Yet, contrary to our expectations, MetaEF training was not more effective than BasicEF training at promoting proactive control. This surprising finding has important implication for understanding proactive control development as it suggests that metacognition may not play a role as critical as previously assumed in that development, at least not in children from 7 years on. It is possible that metacognition is critical for proactive control engagement at ages 5 and 6, when children just start engaging EF proactively but do have much experience yet about task demands where proactive control is adaptive, but not later in development.

An open question is whether any EF training programme can promote proactive control or whether it is necessary to expose children to diversity of tasks tapping different EF processes and gradually increasing in difficulty, as in the present study. This diversity of tasks may provide children the opportunity to test multiple approaches to EF engagement. Recent evidence suggests that acquisition of abstract task knowledge about regularities of a task helps children subsequently engage proactive control (Yanaoka, van't Wout, Saito, & Jarrold, 2022). Giving children the opportunity to experience and learn about different tasks may be sufficient for children to use proactive control adaptively as a function of task demands. This question will have to be tested directly in future studies.

It is worth noting that greater proactive control engagement was observed even though the training task (AX-CPT) involved different task demands from the experimental task (CTS), namely inhibition and task switching, respectively. Although past work has shown that children in middle childhood engage in proactive control consistently across different contexts to a limited extent only (Kubota et al., 2020), our findings show that proactive control can be trained and crucially transferred across untrained contexts in children. Taken together, our findings underline that EF training can help children from low SES backgrounds to engage EF in a more mature and efficient manner, evidencing more flexible EF engagement after EF training.

Change in the way children approached the task was further evidenced by faster responses, at the cost of lower accuracy, from pre- to post-test in the Basic-EF and Meta-EF groups, but not in the control group. Interestingly, the RT reduction was more pronounced for Meta-EF than Basic-EF, perhaps reflecting greater adaptation of strategies in children who received metacognitive training. Even though this pattern was true of both trials with or without early cue presentation, it may be related to greater proactive control engagement given that proactive control can take multiple forms, including not only cue-based preparation but also general attention mobilisation (Jin et al., 2020). On a practical level, one may wonder whether faster responses at the cost of lower accuracy can really be interpreted as effective training benefits rather than a negative outcome. However, encouraging children to shift to greater proactive control strategies often translate into faster responses accompanied by either no gain or even a reduction in accuracy (Chevalier et al., 2015, 2020; Jin et al., 2020). Such process reflects the transient cost of engaging a more mature but less familiar approach (see Siegler & Jenkins, 2014). As such, it should be a transitory phase related to the shift to more proactive EF engagement, and accuracy should increase once children more routinely and efficiently engage EF proactively. However, more research is needed to examine whether and the extent to which the speed accuracy trade-off observed here relates to greater proactive control engagement.

Contrary to our hypotheses, metacognitive training did not yield greater near-transfer effects than the Basic-EF training on any of EF tasks. This is, however, consistent with previous studies that examined the effectiveness of metacognitive training alongside EF training. Both studies (Jones et al., 2020; Pozuelos et al., 2019), in fact, did not find any metacognitive training advantage at the behavioural level at immediate post-test. In a similar vein to our results, Jones et al. found that both groups— working memory training group and metacognitive training group—similarly improved in working memory performance at the immediate post-test, but this effect in a three -month follow-up test in the metacognitive training group only. Pozuelos et al., on the other hand, only found greater metacognitive training effects on a neural marker of conflict processing, but this did not translate to a behavioural advantage. Our findings adds to these other studies that training metacognition and EF in tasks tapping all three EF domains did not yield additional immediate behavioural gains compared with EF training alone. It should be noted, however, that our metacognitive activities emphasised metacognitive monitoring to a greater extent than metacognitive control. As EF has been found to relate more to metacognitive control than monitoring

(Roebers et al., 2012), an EF training emphasising metacognitive control may have yielded different findings.

The metacognitive component of our training did provide children with specific control strategies (e.g., verbal strategies). Instead, it aimed to instill the habit of generating and evaluating control strategies appropriate for new contexts on the basis of task demands and performance monitoring. Thus, the benefits of metacognitive activities may grow over time, as children gain experience with EF engagement in a greater variety of situations, that is, metacognitive training may ignite a positive dynamic of self-reflection on EF engagement. If so, the benefits of instilling metacognitive self-reflection through training may not be fully observable immediately after training and may emerge later after reinforcement has taken place. Indeed, prior studies obtained findings consistent with this possibility. For instance, far-transfer to mathematical skills was not greater after EF training with than without metacognition immediately after training but it was three months later (Jones et al., 2019). Similarly, children who received mindfulness and reflection training (i.e., calming games and activities that encourage reflection in the context of goal-directed problem solving) outperformed the controls only at the follow-up post-test four weeks later (Zelazo et al., 2018). Likewise, the positive effects of 8-week mindfulness training on EF in adolescents with ADHD was strongest at the 8-week follow up test (rather than at immediate post-test or 16-week follow up test; Van de Weijer-Bergsma et al., 2012). In fact, the delayed or gradual manifestation of the effects of intervention—so called the “sleeper effect”—is not uncommon and aligns with the idea that acquired skills such as EF and metacognition need some time for consolidation and independent practice in order to see generalization effects to both near and far skills (Hermida et al., 2015; Zelazo et al., 2018). Thus, future studies should consider running follow-up tests to ensure that there is enough time for the training effects to consolidate into behavioural advantage and to examine whether the observed benefits are maintained for an extended period of time.

Finally, we found no significant interactions between Group (Meta-EF, Basic-EF, Control) and Session (pre, post) for math, reading, and matrix reasoning scores, suggesting that all groups improved their academic performance to a similar degree. The absense of group differences in academic performance gains is partially in line with Jones et al. (2019) in which they found no advantage of metacognitive strategy training over working memory training on both math and reading skills at *immediate post-test* in 9- to 14-year-olds, suggesting that additional metacognitive training did not yield better generalization of far-transfer effects on academic outcomes. Our study, which has similar amount of participants, number of training sessions, length of each training session, and metacognitive training excercises to Jones et al. (2019), also show no effect of EF training, let alone metacognitive and EF training on academic skills. However, Pozuelos et al. (2019) did observe far-transfer to non-verbal reasoning at immediate post-test in 5-year-olds, perhaps because far-transfer is greater in younger children.

Multiple meta-analyses showed weak to modest far-transfer effects of process-based EF training on academic skills (Melby-Lervåg et al., 2016; Sala & Gobet, 2017). A large scale study investigating training effects of core modules of EF (updating switching, and inhibition) showed no evidence of transfer effects to non-trained EF tasks and training-specific effects were observed for updating only (Zuber et al., 2023). Some researchers even suggest that current theoretical framework that reduces EF to a few component processes that support other abilities/phenomena or self regulation should be reconsidered, as how we conceptualize EF has a great impact on how we construct and implement EF training and interventions (Doebel et al., 2017; Perone et al., 2021). Our findings, thus, suggest that added benefits of metacognitive training in addition to EF training is mininmal for academic skills, although null effects should not be interpreted as an *absence* of an effect. Since there are limited amount of work that examine the (a) training effects of metacognition on EF, as well as (b) EF and metacognitive training on EF, the results so far are mixed and requires further investigation (and perhaps reframing/reconsideration of training programmes) with large sample size.

## 5. Conclusion

Both metacognitive training and basic EF training yielded greater working memory and proactive control improvement relative to the controls. Thus, proactive control engagement is amenable to change after cognitive training in children, even though EF training did not target proactive control specifically. Further, adding metacognition to EF training did not yield greater increase in proactive control engagement. These findings have important conceptual implication. First, they suggest that proactive control engagement is malleable through training in children. Second, they failed to provide support for the role of metacognition in proactive control development, either because metacognitive training needs to focus more explicitly on reactive/proactive control or because reactive/proactive control adjustments are to a large extent implicit (Gonthier, Ambrosi, & Blaye, 2021). Third, they show that EF training programmes can induce both qualitative and quantitative changes in the ways in which children engage EF processes. However, we found no evidence for EF alone or EF and metacognition training benefits on academic skills. Our findings thus suggest minimal added benefits of metacognitive training (along with EF training), as no greater near- or far-transfer effects were observed in the metacognitive training group. By demonstrating the potential benefits of EF training, with or without metacognitive reflection training for children from low SES backgrounds, we hope this work will help generate concrete educational instruments to counter socioeconomic disparities and improve children's opportunities for educational equality.

## References

Andersen, L. M., Visser, I., Crone, E. A., Koolschijn, P., & Raijmakers, M. E. (2014). Cognitive strategy use as an index of developmental differences in neural responses to feedback. *Developmental Psychology*, *50*(12), 2686.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bergman Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science*, *14*(3), 591–601.

Bigorra, A., Garolera, M., Guijarro, S., & Hervás, A. (2016). Long-term far-transfer effects of working memory training in children with ADHD: A randomized controlled trial. *European Child & Adolescent Psychiatry*, *25*(8), 853–867.

Blackwell, K. A., Chatham, C. H., Wiseheart, M., & Munakata, Y. (2014). A developmental window into trade-offs in executive function: The case of task switching versus response inhibition in 6-year-olds. *Neuropsychologia*, *62*, 356–364.

Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science*, *17*(2), 203–211.

Blakey, E., & Carroll, D. J. (2015). A short executive function training program improves preschoolers' working memory. *Frontiers in Psychology*, *6*, 1827.

Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, *43*(6), 541–552. https://doi.org/10.1177/0022219410371676

Brown, A. L. (2017). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension: Perspectives* *from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education* (pp. 453–482). Routledge.

Bryce, D., Whitebread, D., & Szucs, D. (2015). The relationships among executive functions, metacognitive skills and educational achievement in 5 and 7 year-old children. *Metacognition and Learning*, *10*, 181–198.

Bryce, D., & Whitebread (2012). The development of metacognitive skills: evidence from observational analysis of young children's behavior during problem-solving. Metacognition Learning, 7, 197-217.DOI 10.1007/s11409-012-9091-2

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404.

Carretti, B., Borella, E., Elosúa, M. R., Gómez-Veiga, I., & García-Madruga, J. A. (2017). Improvements in reading comprehension performance after a training program focusing on executive processes of working memory. *Journal of Cognitive Enhancement*, *1*(3), 268–279.

Chacko, A., Bedard, A. C., Marks, D. J., Feirsen, N., Uderman, J. Z., Chimiklis, A., Rajwan, E., Cornwell, M., Anderson, L., & Zwilling, A. (2014). A randomized clinical trial of Cogmed working memory training in school-age children with ADHD: A replication in a diverse sample using a control condition. *Journal of Child Psychology and Psychiatry*, *55*(3), 247–255.

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*(14), 5529–5533. https://doi.org/10.1073/pnas.0810002106

Chen, Z., & Siegler, R. S. (2000). VII. Components of Strategic Change. *Monographs of the Society for Research in Child Development*, *65*(2), 43–58. https://doi.org/10.1111/1540-5834.00080

Chevalier, N. (2015). The development of executive function: Toward more optimal coordination of control with age. *Child Development Perspectives*, *9*(4), 239–244.

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. https://doi.org/10.1037/a0019674

Chevalier, N., & Blaye, A. (2016). Metacognitive monitoring of executive control engagement during childhood. *Child Development,* 87 (4), 1264–1276. https://doi.org/ 10.1111/cdev.12537.

Chevalier, N., James, T. D., Wiebe, S. A., Nelson, J. M., & Espy, K. A. (2014). Contribution of reactive and proactive control to children's working memory performance: Insight from item recall durations in response sequence planning. *Developmental Psychology*, *50*(7), 1999.

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience*, *27*(6), 1125–1136.

Chevalier, N., Meaney, J. A., Traut, H. J., & Munakata, Y. (2020). Adaptiveness in proactive control engagement in children and adults. *Developmental Cognitive Neuroscience*, *46*, 100870.

Cragg, L., & Nation, K. (2010). Language and the Development of Cognitive Control. *Topics in Cognitive Science*, *2*(4), 631–642. https://doi.org/10.1111/j.1756-8765.2009.01080.x

Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5- to 7-year-old children. *Journal of Experimental Child Psychology*, *126*, 213–228. https://doi.org/10.1016/j.jecp.2014.04.001

Diamond, A. (2012). Activities and programs that improve children’s executive functions. *Current Directions in Psychological Science*, *21*(5), 335–341.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science (New York, N.Y.)*, *333*(6045), 959–964. https://doi.org/10.1126/science.1204529

Doebel, S., Barker, J. E., Chevalier, N., Michaelson, L. E., Fisher, A. V., & Munakata, Y. (2017). Getting ready to use control: Advances in the measurement of young children’s use of proactive control. *PloS One*, *12*(4), e0175072.

Dunning, D. L., Holmes, J., & Gathercole, S. E. (2013). Does working memory training lead to generalized improvements in children with low working memory? A randomized controlled trial. *Developmental Science*, *16*(6), 915–925.

DuPuis, D., Ram, N., Willner, C. J., Karalunas, S., Segalowitz, S. J., & Gatzke-Kopp, L. M. (2015). Implications of ongoing neural development for the measurement of the error-related negativity in childhood. *Developmental Science*, *18*(3), 452–468.

Espinet, S. D., Anderson, J. E., & Zelazo, P. D. (2012). N2 amplitude as a neural marker of executive function in young children: An ERP study of children who switch versus perseverate on the Dimensional Change Card Sort. *Developmental Cognitive Neuroscience*, *2*, S49–S58.

Fatzer, S. T., & Roebers, C. M. (2012). Language and Executive Functions: The Effect of Articulatory Suppression on Executive Functioning in Children. *Journal of Cognition and Development*, *13*(4), 454–472. https://doi.org/10.1080/15248372.2011.608322

Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., Locke, J., Ishijima, E., & Kasari, C. (2010). Effects of mindful awareness practices on executive functions in elementary school children. *Journal of Applied School Psychology*.

Ganasan, K., & Steinbeis, N. (2021). Effort-related decision-making and its underlying processes during childhood. *Developmental Psychology*, 57(9), 1487-1496. https://doi.org/10.1037/dev0001228

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.

Gonthier, C., Ambrosi, S., & Blaye, A. (2021). Learning-based before intentional cognitive control: Developmental evidence for a dissociation between implicit and explicit control. Journal of Experimental Psychology: Learning,Memory, and Cognition, https://doi.org/10.1037/xlm0001005

Hadley, L. V., Acluche, F., & Chevalier, N. (2019). Encouraging performance monitoring promotes proactive control in children. *Developmental Science*, e12861.

Haugen, I., Ueland, T., Stubberud, J., Brunborg, C., Wykes, T., Øie, M. G., & Haug, E. (2023). Moderators of metacognitive strategy training for executive functioning in early schizophrenia and psychosis risk. *Schizophrenia Research: Cognition*, *31*, 100275.

Hermida, M. J., Segretin, M. S., Prats, L. M., Fracchia, C. S., Colombo, J. A., & Lipina, S. J. (2015). Cognitive neuroscience, developmental psychology, and education: Interdisciplinary development of an intervention for low socioeconomic status kindergarten children. *Trends in Neuroscience and Education*, *4*(1–2), 15–25.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), F9–F15.

Hovik, K. T., Saunes, B.-K., Aarlien, A. K., & Egeland, J. (2013). RCT of working memory training in ADHD: Long-term near-transfer effects. *PLoS One*, *8*(12), e80561.

Jin, X., Auyeung, B., & Chevalier, N. (2020). External rewards and positive stimuli promote different cognitive control engagement strategies in children. *Developmental Cognitive Neuroscience*, *44*, 100806.

Jones, B. K., Destin, M., & McAdams, D. P. (2018). Telling better stories: Competence-building narrative themes increase adolescent persistence and academic achievement. *Journal of Experimental Social Psychology*, *76*, 76–80.

Jones, J. S., Milton, F., Mostazir, M., & Adlam, A. R. (2020). The academic outcomes of working memory and metacognitive strategy training in children: A double-blind randomized controlled trial. *Developmental Science*, *23*(4), e12870.

Kajka, N. (2019). The influence of metacognitive training on the improvement of working memory in children with ADHD. *Current Problems of Psychiatry*, *20*(3), 217–227.

Kälin, S., & Roebers, C. M. (2022). Longitudinal associations between executive functions and metacognitive monitoring in 5-to 8-year-olds. *Metacognition and Learning*, 1–17.

Karbach, J., Könen, T., & Spengler, M. (2017). Who benefits the most? Individual differences in the transfer of executive control training across the lifespan. *Journal of Cognitive Enhancement*, *1*(4), 394–405.

Karbach, J., Strobach, T., & Schubert, T. (2015). Adaptive working-memory training benefits reading, but not mathematics in middle childhood. *Child Neuropsychology*, *21*(3), 285–301.

Karbach, J., & Unger, K. (2014). Executive control training from middle childhood to adolescence. *Frontiers in Psychology*, *5*(May), 390. https://doi.org/10.3389/fpsyg.2014.00390

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147.

Kassai, R., Futo, J., Demetrovics, Z., & Takacs, Z. K. (2019). A meta-analysis of the experimental evidence on the near-and far-transfer effects among children's executive function skills. *Psychological Bulletin*, *145*(2), 165.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD-a randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(2), 177–186.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology. General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

Kray, J., Karbach, J., Haenig, S., & Freitag, C. (2012). Can task-switching training enhance executive control functioning in children with attention deficit/-hyperactivity disorder? *Frontiers in Human Neuroscience*, *5*, 180.

Krebs, S. S., & Roebers, C. M. (2010). Children's strategic regulation, metacognitive monitoring, and control processes during test taking. *British Journal ofEducational Psychology*, 80, 325–340.

Kubota, M., Hadley, L. V., Schaeffner, S., Könen, T., Meaney, J.-A., Auyeung, B., Morey, C. C., Karbach, J., & Chevalier, N. (2020). Consistent use of proactive control and relation with academic achievement in childhood. *Cognition*, *203*, 104329.

Kuhn, D. (2000). Metacognitive Development. *Current Directions in Psychological Science*, *9*(5), 178–181. https://doi.org/10.1111/1467-8721.00088

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, *84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096

Lemaire, P., & Brun, F. (2013). Effects of strategy sequences and response-stimulus intervals on children's strategy selection and strategy execution: A study in computational estimation. *Psychological Research*. https://doi.org/10.1007/s00426-013-0501-0

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* (R package version 1.6.2-1.).

Lucenet, J., & Blaye, A. (2014). Age-related changes in the temporal dynamics of executive control: A study in 5- and 6-year-old children. *Frontiers in Psychology*, *5*(July), 1–11. https://doi.org/10.3389/fpsyg.2014.00831

Lyons, K. E., & Zelazo, P. D. (2011). Monitoring, metacognition, and executive function: Elucidating the role of self-reflection in the development of selfregulation. In B. B. Janette (Ed.). Advances in child development and behavior (Vol. 40, pp. 379–412). Oxford, UK: Elsevier.

Marulis, L. M., Baker, S. T., & Whitebread, D. (2020). Integrating metacognition and executive function to enhance young children's perception of and agency in their learning. *Early Childhood Research Quarterly*, *50*, 46–54.

McGuire, J. T., & Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(17), 7922–7926. https://doi.org/10.1073/pnas.0910662107

Melby-Lervåag, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer" evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*(4), 512–534.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Moriguchi, Y., Sakata, Y., Ishibashi, M., & Ishikawa, Y. (2015). Teaching others rule-use improves executive function and prefrontal activations in young children. *Frontiers in Psychology*, *6*, 894.

Munakata, Y. E., Snyder, H. R., & Chatham, C. H. (2012).Developing cognitive control: Three key transitions.*Current Directions in Psychological Science*, 21, 71–77. doi:10.1177/0963721412436807

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. The Psychology of Learning and Motivation, 26, 125–173.

Niebaum, J. C., Chevalier, N., Guild, R. M., & Munakata, Y. (2019). Adaptive control and the avoidance of cognitive control demands across development. *Neuropsychologia*, 123, 152–158. https://doi.org/10.1016/ j.neuropsychologia.2018.04.029

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, *13*(2), 179–212.

O'Leary, A.P., Sloutsky, V.M., 2017. Carving metacognition at its joints: protracted development of component processes. *Child Development*, 88 (3), 1015–1032.

Partanen, P., Jansson, B., Lisspers, J., & Sundin, Ö. (2015). Metacognitive strategy training adds to the effects of working memory training in children with special educational needs. *International Journal of Psychological Studies*, *7*(3), 130–140.

Pereg, M., Shahar, N., & Meiran, N. (2013). Task switching training effects are mediated by working-memory management. *Intelligence*, *41*(5), 467–478.

Perone, S., Simmering, V. R., & Buss, A. T. (2021). A dynamical reconceptualization of executive-function development. *Perspectives on Psychological Science*, *16*(6), 1198–1208.

Pozuelos, J. P., Combita, L. M., Abundis, A., Paz-Alonso, P. M., Conejero, Á., Guerra, S., & Rueda, M. R. (2019). Metacognitive scaffolding boosts cognitive and neural benefits following executive attention training in children. *Developmental Science*, *22*(2), e12756.

Re, A. M., Capodieci, A., & Cornoldi, C. (2015). Effect of training focused on executive functions (attention, inhibition, and working memory) in preschoolers exhibiting ADHD symptoms. *Frontiers in Psychology*, *6*, 1161.

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review*, *45*, 31–51.

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. Metacognition and Learning, 7, 151–173.

Roebers, C. M., & Feurer, E. (2016). Linking executive functions and procedural metacognition. *Child Development Perspectives*, *10*(1), 39–44.

Roebers, C. M., Krebs, S. S., & Roderer, T. (2014). Metacognitive monitoring and control in elementary school children: Their interrelations and their role for test performance.

*Learning and Individual Differences*, *29*, 141–149. https://doi.org/10.1016/j.lindif.2012.12.003

Röthlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., & Roebers, C. M. (2012). Improving executive functions in 5-and 6-year-olds: Evaluation of a small group intervention in prekindergarten and kindergarten children. *Infant and Child Development*, *21*(4), 411–429.

Sala, G., & Gobet, F. (2017). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*(4), 671.

Schaeffner, S., Chevalier, N., Kubota, M., & Karbach, J. (2021). Metacognitive training. In *Cognitive Training* (pp. 255–270). Springer.

Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain, and Education*, *2*(3), 114–121.

Schonert-Reichl, K. A., & Lawlor, M. S. (2010). The effects of a mindfulness-based education program on pre-and early adolescents' well-being and social and emotional competence. *Mindfulness*, *1*, 137–151.

Segretin, M. S., Lipina, S. J., Hermida, M. J., Sheffield, T. D., Nelson, J. M., Espy, K. a, & Colombo, J. a. (2014). Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds. *Frontiers in Psychology*, *5*(March), 205. https://doi.org/10.3389/fpsyg.2014.00205

Shapiro, S. L., Lyons, K. E., Miller, R. C., Butler, B., Vieten, C., & Zelazo, P. D. (2015). Contemplation in the classroom: A new direction for improving childhood education. *Educational Psychology Review, 27*, 1–30.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Siegler, R., & Jenkins, E. A. (2014). *How children discover new strategies*. Psychology Press.

Spiess, M. A., Meier, B., & Roebers, C. M. (2016). Development and longitudinal relationships between children's executive functions, prospective memory,and metacognition. *Cognitive Development*, 38, 99–113.

Skau, S., Helenius, O., Sundberg, K., Bunketorp-Käll, L., & Kuhn, H. G. (2022). Proactive cognitive control, mathematical cognition and functional activity in the frontal and parietal cortex in primary school children: An fNIRS study. *Trends in Neuroscience and Education*, 28, 100180.

Strobach, T., & Karbach, J. (2021). *Cognitive Training—An Overview of Features and Application* (2nd ed.). Springer.

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106–113. https://doi.org/10.1111/j.1467-7687.2008.00745.x

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, *78*(6), 852–868.

Traverso, L., Viterbori, P., & Usai, M. C. (2015). Improving executive function in childhood: Evaluation of a training intervention for 5-year-old children. *Frontiers in Psychology*, *6*, 525.

Van de Weijer-Bergsma, E., Formsma, A. R., de Bruin, E. I., & Bögels, S. M. (2012). The effectiveness of mindfulness training on behavioral problems and attentional functioning in adolescents with ADHD. *Journal of Child and Family Studies*, *21*(5), 775–787.

Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*(1), 3–14.

Voigt, B., Mahy, C. E., Ellis, J., Schnitzspahn, K., Krause, I., Altgassen, M., & Kliegel, M. (2014). The development of time-based prospective memory in childhood: The role of working memory updating. *Developmental Psychology*, *50*(10), 2393.

Wang, C., Li, B., & Yao, Y. (2021) Proactive Control Mediates the Relationship Between Working Memory and Math Ability in Early Childhood. *Front. Psychol.* 12:611429. doi: 10.3389/fpsyg.2021.611429

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children* (Fourth edition). Pearson Education.

Wechsler, D. (2017). *Wechsler Individual Achievement Test Third Edition* (Third edition). Pearson Education.

Wiemers, E. A., & Redick, T. S. (2018). Working memory capacity and intra-individual variability of proactive control. *Acta Psychologica*, *182*, 21–31.

Wong, A. S., He, M. Y., & Chan, R. W. (2014). Effectiveness of computerized working memory training program in Chinese community settings for children with poor working memory. *Journal of Attention Disorders*, *18*(4), 318–330.

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences*, *8*(1), 12–17. https://doi.org/10.1016/j.tics.2003.11.001

Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. *Developmental Review*, 38, 55–68.

Zelazo, P. D., Forston, J. L., Masten, A. S., & Carlson, S. M. (2018). Mindfulness plus reflection training: Effects on executive function in early childhood. *Frontiers in Psychology*, *9*, 208.

Zuber, S., Joly-Burra, E., Mahy, C. E., Loaiza, V., & Kliegel, M. (2023). Are facet-specific task trainings efficient in improving children's executive functions and why (they might not be)? A multi-facet latent change score approach. *Journal of Experimental Child Psychology*, *227*, 105602.