Title

The malleability of executive function in early childhood: Effects of schooling and targeted training

Journal

Developmental Science, 22(2)

ISSN
1363-755X

Authors

Zhang, Qiong
Wang, Cuiping
Zhao, Qianwen
et al.

Publication Date
2019-03-01

DOI
10.1111/desc.12748

Published in final edited form as:

Dev Sci. 2019 March ; 22(2): e12748. doi:10.1111/desc.12748.

The malleability of executive function in early childhood: effects of schooling and targeted training

Qiong Zhang[1], **Cuiping Wang**[1], **Qianwen Zhao**[2], **Ling Yang**[1], **Martin Buschkuehl**[3], and **Susanne M. Jaeggi**[4]

[1]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China

[2]Hangzhou No. 4 High School, Hangzhou, China

[3]MIND Research Institute, Irvine, CA, USA

[4]School of Education, University of California, Irvine, Irvine, USA

Abstract

Executive function (EF), its importance for scholastic achievement and the question of whether or not EF is malleable have become a topic of intense interest. Education or schooling is often seen as effective approaches to enhance EF due to the specific school-related requirements as compared to kindergarten or pre-school. However, no study to date has investigated whether targeted training focusing on those domains might be comparable with regular schooling in improving EF and fluid intelligence (Gf). The aim of the present study was to replicate and extend the previously demonstrated schooling effects on EF by using a school-cutoff design, and to further investigate whether a theoretically motivated intervention targeting specific EF, i.e., working memory (WM) or inhibitory control (IC), could achieve comparable effects with schooling in both, WM and IC, as well as Gf. 91 six-year-old kindergarteners and first-graders with similar chronological age participated the study. We compared the performance of a first-grade schooling group with that of two kindergarten training groups as well as a business-as-usual kindergarten control group. Participants were assessed in WM, IC and Gf at baseline, immediately after the intervention (posttest), as well as 3 months after training completion (follow-up). The results showed that the schooling group indeed outperformed the kindergarten groups at baseline in several cognitive tasks. Furthermore, both the WM and IC training showed pronounced gains in the trained tasks, as well as varying degrees of improvement in non-trained outcome measures. Most importantly, both training groups achieved comparable performance with the schooling group, which was especially apparent in Gf at follow-up. Our findings provide further evidence for the malleability of EF demonstrating that both, long-term and short-term interventions can facilitate the acquisition of those important skills, and as such, our work has important implications for educational practice.

Corresponding author: Susanne M. Jaeggi, 3452 Education, University of California, Irvine, Irvine, CA 92697-5500, smjaeggi@uci.edu, Phone: 949-824-5896.

Author note: M.B.is employed at the MIND Research Institute, whose interest is related to this work, and S.M.J. has an indirect financial interest in MIND Research Institute.

Keywords

Cognitive training; Transfer; Working memory; Inhibitory control; Fluid intelligence; Education

Executive function (EF) refers to a group of cognitive processes that are required when individuals have to engage in goal-directed behavior (Zelazo & Carlson, 2012). There is general agreement that EF includes a number of interrelated processes, such as working memory (WM), inhibitory control (IC) and cognitive flexibility (Miyake et al., 2000). It has been argued that WM and IC are two of the most fundamental functions underlying higher cognitive functions (Christoff, Ream, Geddes, & Gabrieli, 2003; Diamond, 2013).

EF and the supporting brain structures, especially the prefrontal cortex develop rapidly in early childhood and mature well into late adolescence (Carlson, Zelazo, & Faja, 2013; Diamond, 2013). In children, EF is closely related to scholastic achievement (e.g., Alloway et al., 2005; Alloway & Alloway, 2010) and school-related behaviors (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Gathercole, Lamont, & Alloway, 2006). Deficits in EF are thought to be an important cause of cognitive impairment in children with attention deficit hyperactivity disorder (ADHD; Alderson, Rapport, & Kofler, 2007), and other learning disabilities (Minear & Shah, 2006).

Given the importance of EF, the question of whether and how we might improve EF has been a hot topic. Researchers have suggested that education and schooling affect performance in cognitive tasks assessing EF (e.g., Baker et al., 2015; Brod, Bunge, & Shing, 2017; Ceci, 1991). In addition, studies have provided evidence that the development of EF might be fostered by specific school-related requirements in early primary school children (e.g., Blair & Raver, 2014; Burrage et al., 2008; Diamond, Barnett, Thomas, & Munro, 2007; Raver et al., 2008, 2011).

But EF can not only be improved through schooling. Accumulating evidence showed that targeted cognitive interventions can be effective as well. Most of these training studies have been focusing on WM. For example, it has been widely demonstrated that training skills related to WM does not only improve performance in the trained tasks, but also generalize to untrained tasks, such as reading comprehension (e.g., Chein & Morrison, 2010), or fluid reasoning (e.g., Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Klingberg et al., 2005; Schmiedek, Lövdén, & Lindenberger, 2010). Despite these encouraging findings, such transfer effects are often small and not consistent across studies (cf. Au et al., 2016, 2015; Karbach, & Verhaeghen, 2014; Melby-Lervåg & Hulme, 2013; Melby-Lervåg, Redick, & Hulme, 2016; Soveri, Antfolk, Karlsson, Salo, & Laine, 2017; Weicker, Villringer, & Thöne-Otto, 2016 for recent meta-analyses), and importantly, there are only very few studies to date attempting to implement cognitive training in early childhood (Kroesbergen, Van't Noordende, & Kolkman, 2012; Ramani, Jaeggi, Daubert, & Buschkuehl, 2017).

Furthermore, there are only very few studies that focus on training targeted IC, which develops markedly in early childhood and is also a key predictor of school readiness for preschoolers and school achievement in mathematics and reading (Blair & Razza, 2007; Davidson, Amso, Anderson, & Diamond, 2006; Willoughby, Kupersmidt, & Voegler-Lee, 2012). Although there are some curriculum-based interventions (e.g., Tools of the Mind, e.g., Blair & Raver, 2014; Diamond et al., 2007) and game-like training programs (Wexler et al., 2016) that include aspects of IC which have been successful, there are almost no studies to date that used computerized interventions in early childhood, and furthermore, the reported training and transfer effects have been inconsistent. For instance, Rueda et al. (2005) trained 4- and 6-year-old children with an attention training battery that included a Stroop-like task. Transfer effect was assessed with a flanker task, the Kaufman Brief Intelligence Test, and the Children's Behavior Questionnaire. Significant transfer effects were observed in the Kaufman Brief Intelligence Test, and although there were training-related differences in the brain activation pattern during the flanker task, there were no behavioral effects. Similarly, Rueda, Checa, and Combita (2012) trained 5-year-old children with the same training program that included a Stroop-like task. Their results also showed transfer effects to fluid intelligence (Gf), and the brain circuitry involved in executive attention was activated faster and more efficiently after training. Another study used interventions focusing on WM and IC (go/no-go, flanker, stop-signal task) in 4- to 5-year-old children (Thorell et al., 2009). They only found the IC training effect in the go/no-go and flanker tasks, but no transfer effects. On the other hand, the WM training resulted not only in improved WM performance, but also in transfer to IC.

Overall, despite the mostly small and inconsistent effects, there is some evidence that both, schooling and targeted training might lead to changes in EF, providing evidence for the malleability of EF, especially in young children (Wass, Scerif, & Jonhnson, 2012). However, to the best of our knowledge, there is no study to date that has systematically compared the effects of targeted cognitive training with the effects of schooling. The goal of the present study was to add to that literature and to explore whether targeted training might lead to similar cognitive changes that are observed as a result of regular schooling. Specifically, our study had two main goals: first, we aimed at replicating and extending the previously demonstrated schooling effects in various measures of EF and Gf. Second, we tested the effects of targeted training on EF in kindergartners (by focusing either on WM or on IC skills) and investigated whether the potential improvements in the trained groups would be comparable with those emerging in a schooling group with similar chronological age.

To achieve our goals, we implemented a school-cutoff design, which allowed us to tease apart schooling effects from other non-school-related factors (Burrage et al., 2008; Morrison, Smith, & Dow-Ehrensberger, 1995). For example, if the local cutoff date for primary school is September 1st, all children reaching age 6 before September 1st will begin elementary school, while those born after September 1st will enroll in the following year and will attend kindergarten instead. As such, using this school-cutoff design will allow us to compare participants with similar chronological age who differ in their educational experiences. Specifically, we can test the schooling effect by comparing the cognitive performance of children in first grade with that of their age-matched kindergarten peers. Likewise, comparing the performance of the WM and the IC training groups with those of the kindergarten and the schooling groups allows us to investigate whether targeted training improves EF, and if so, whether those improvements would be comparable with those observed as a function of schooling. In other words, we can test whether targeted training would enable kindergarten groups to catch up with the schooling group.

Given the critical role of WM and IC in Gf, and because of previous studies demonstrating improvements in Gf after training EF (e.g., Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Rueda et al., 2012), we also included Gf as a dependent variable. Finally, given that earlier work with children has shown that training-related gains can persist for three months (Jaeggi et al., 2011; Klingberg, Forssberg, & Westerberg, 2002; Karbach, Strobach, & Schubert, 2015), we also tested for any longitudinal effects of training three months after the posttest.

Methods

Participants

Based on the results of our pilot studies, we predicted a medium effect size (Cohen's f = 0.20; Cohen, 1988) for our experimental design. A power analysis using G*power 3.1 determined that given an α level of 0.05 and a power of 0.80, the sample size required to achieve the predicted effect size was approximately 75 individuals.

Ninety-one typically-developing children (25 "young" first graders and 66 "old" kindergarteners; mean age: 73.50 months; SD: 2.10) from a primary school and two kindergartens in Hangzhou (China) participated the study. The first graders (11 girls) who participated in regular school classes and activities made up the schooling group (SG) whose birthdates fell within three months before the cutoff date of September 1st (from June 1st to August 31st), which was set by local authorities as the entry date for primary school. The birth dates of the kindergarteners fell within three months after the cutoff date (from September 1st to November 30th). Among the kindergarteners, 22 children (10 girls) were assigned to the business-as-usual kindergarten control group (CG). The remaining 44 children were assigned to either the WM training group or the IC training group. Two children withdrew from the study after the pretest (one from the WM group, and one from the IC group). Another child from the WM group dropped out after the first training session. Thus, the data of 20 children (10 girls) in the WM training group, and 21 children (10 girls) in the IC training group were included in the final data analysis. All participants were from upper-middle-income families. The four groups were matched according to gender (BF_{10} = 0.05), and in addition, the three kindergarten groups were matched with respect to age (BF_{10} = 0.36) and Gf scores (BF_{10} = 0.15). The study was approved by the ethical committee at Zhejiang University (China), and written consent from caregivers and teachers were obtained for all participants.

Procedure

We implemented a 4 (groups: SG/ WM training group/ IC training group/ CG) × 3 (session: pretest/ posttest/ follow-up test) experimental design. During the intervention period, the first graders and kindergartners of the control group took part in their regular curricular activities. The two kindergarten training groups completed a 15-minute intervention each day, 4 times per week, for a total of 20 sessions in addition to their kindergarten routine activities. All participants completed three assessments, i.e., the pretest, the posttest, and the follow-up test. The pretest and the posttest were conducted within a week before and after the intervention period. The follow-up test was completed approximately three months after the posttest.

Interventions

Working memory training—The WM training procedure has been developed and used in previous research (Loosli, Buschkuehl, Perrig, & Jaeggi, 2012; see Figure 1a). The task consisted of two parts: in the encoding stage, a sequence of animals was presented in the center of the screen. The participants were asked to identify the orientation of the picture by pressing an appropriate button as soon as an animal appeared. At the same time, they were required to remember the order in which the animals were presented. If children failed to provide a response within 3000ms or hit the wrong button, an error was indicated. In the recall stage, children had to reproduce the previously shown animal sequence in order without time limit. At the end of each trial, feedback was provided. In addition, a performance indicator was given as a high score representing the maximum number of animals that could be reproduced in the correct order providing that there was no mistake in the orientation identification part. The difficulty level of the task was continuously adapted, i.e., set size increased or decreased depending on the performance in the previous trial. Each game started with a set size of two (i.e., two animals) and ended after approximately 6 minutes. In each training session, children played the game twice for about 15 minutes. The maximum set size that the children achieved per training session (i.e. their WM span) served as dependent variable.

Inhibitory control training—Two child-friendly tasks were used for the IC intervention: An adaptive stop-signal task (Berkman, Kahn, & Merchant, 2014), and a modified Stroop task (Rueda et al., 2012).

Stop-signal task.: The go stimuli comprised of a strawberry and a peach (150×150 pixels in size) (see Figure 1b). Participants were instructed to press “F” when a strawberry appeared and “J” when a peach did (go-trials), but they had to inhibit their response when an apple (150×150 pixels in size) appeared shortly after the presentation of the go-stimulus (stop-trials). Each trial began with a central fixation lasting for 500ms, followed by a go-stimulus with a 1,500ms duration in the case of go-trials. The stop-signal delay (SSD), the time between go- and stop-signals in stop-trials, was adjusted individually for each participant by using a staircase function. The first stop trial started with a 250ms delay, and the subsequent trials were adjusted by adding or subtracting 50ms in the case of successful or failed inhibition, respectively (Berkman et al., 2014). Lower and upper boundaries for the SSD tracking procedure were set to 100ms and 400ms, respectively, as determined by pilot experiments. There were 80 trials per session, of which 25% were stop-trials, randomly presented throughout the session. Each session lasted approximately 8 minutes. Performance feedback was provided at the end of each session. As dependent variable, we used stop-signal reaction time (SSRT), which was calculated by subtracting the mean SSD from the $\lceil n \rceil$th RT (the integration method, Verbruggen, Chambers, & Logan, 2013).

Modified Stroop Task.: The conflicting dimensions were number and size (see Figure 1c). The task included two versions with different stimuli (either fruits or animals). In each trial, participants were presented with various numbers of cherries (or mice) and watermelons (or elephants) at the same time, and they were asked to indicate which set had more than the other, irrespective of fruit/animal size. There were 100 trials per task variant, of which 50%

congruent trials (where the larger number of stimuli was represented with the larger stimuli, e.g., two cherries vs. five watermelons) and 50% incongruent trials (e.g., five cherries vs. two watermelons). The congruent trials and incongruent trials were randomly intermixed. Each training session lasted about 5 minutes. All children completed the fruit version first, and then the animal version. Performance feedback was provided at the end of each session. The response accuracy of both, incongruent and congruent trials (in percent), reaction times for both trial types (correct responses only) served as the dependent measures.

Outcome Measures

Backward Digit Span Task **(BDST; Prencipe et al., 2011; Wechsler, 1991).—**

Children were instructed to listen to a sequence of single digit numbers and then to repeat those numbers in the reverse order in which they were presented. The lists of numbers started with two digits and each difficulty level appeared twice. Testing was discontinued after the child responded incorrectly on both trials at a certain level. The total amount of trials participants had completed correctly was used as dependent variable.

AX-CPT **(Braver, Satpute, Rush, Racine, & Barch, 2005).—**To make the task more child-friendly, all the letter stimuli in the adult version were replaced by ubiquitous animal pictures. A target response was required when an A cue (panda) was followed by an X probe (giraffe), whereas non-target responses were to be given for all other cue–probe pairs (AY, BX, and BY trials, where Y and B represent any animals other than panda and giraffe). AX trials made up 70% of trials, while the frequency of each of the other three types of non-target trials was 10%. The task consisted of 160 trials, lasting about 9 minutes. Given that we were interested in changes in IC as a function of training, we focused particularly on the proportion of errors and correct RT in AY and BX type trials, which reflect reactive and proactive control adjustments (Braver, 2012). Thus, we used the Behavioral Shift Index (BSI) as the dependent variable (Braver, Paxton, Locke, & Barch, 2009; Chiew & Braver, 2014; Maraver, Bajo, & Gomez-Ariza, 2016)[1]. The BSI is based on the formula $(AY- BX)/(AY+ BX)$ for errors and correct RT. Trials where errors were equal to 0 were corrected to $(\text{errors} + 0.5)/(\text{frequency of trials} + 1)$. The BSI calculation yields a score between −1 and +1: The closer a score is to +1, the more proactive a participant's strategy; the closer a score is to −1, the more reactive a participant's strategy. Using BSI as dependent variable allows us to illustrate a potential maturational shift from using predominantly reactive strategies to relying on more proactive control processes, which has been suggested to happen between the ages 5 and 7 (see Braver, 2012; Chevalier, James, Wiebe, Nelson, & Espy, 2014; Lucenet & Blaye, 2014).

Raven's Standard Progressive Matrices **(RSPM, Raven, Court, & Raven, 1998).—**Participants were asked to choose the appropriate pattern from a set of given patterns to logically complete a given matrix. We split the RSPM into odd and even items (30 items per version, including one practice item per version), and the versions A and B were administered in counterbalanced order across the three sessions (i.e. either ABA or BAB[2]).

---

[1] Note, we also report the descriptive data for the individual AX-CPT trials (AX, AY, BX, BY) in the supplementary material (cf. Table S1).

Participants were given 20 minutes to complete the test, and the dependent variable was the number of correct responses provided in that time.

The training tasks and the computerized assessment task (i.e. BDST, AX-CPT) were programmed in E-prime 1.1, and children were tested with 14-inch laptops in a one-on-one setting. All tasks were presented in the same order, that is, the children completed BDST, AX-CPT and then RSPM.

Analytical Approach

To test for baseline differences, we ran two separate univariate analyses of variance (ANOVA) for each of the outcome variables. The first one tested whether the three kindergarten groups differed from each other verifying whether they were equally matched. The second ANOVA aimed at specifically testing the schooling effect by combining the kindergarten groups into one group (providing that the kindergarten groups did not differ from each other), and then comparing their performance with that of the first-graders (SG), hypothesizing that the first-graders would outperform the kindergartners at baseline.

To test for specific training effects, we used paired t-tests to examine the improvement over time on the trained tasks by comparing the average performance of the first week (averaged across 4 sessions) with the average performance of the last week (averaged across 4 sessions) in each of the trained tasks.

Given our particular design and our study aim that uses intact groups (first graders vs. kindergartners), we conducted a set of mixed ANOVAs for each cognitive measure to test for changes as a function of training and/or schooling. Specifically, to test for immediate changes, we conducted mixed ANOVAs using ‘Session’ (Pre, Post) as within-group factor and ‘Group’ (WM, IC, CG, SG) as between-group factor. Similarly, to investigate longitudinal changes, we calculated mixed ANOVAs using ‘Session’ (Pre, Follow-up) as within-group factor and ‘Group’ (WM, IC, CG, SG) as between-group factor. To follow up those analyses, we also calculated paired comparisons to specifically investigate the change from pretest to posttest and from pretest to follow-up test for each group, and furthermore, to assess transfer, we calculated mixed ANOVAs using ‘Session’ (Pre vs. Post, or Pre vs. Follow-up, respectively) as within-group factor, and ‘Group’ (CG vs. WM or CG vs. IC, respectively) as between-group factor. Finally, to test whether a) schooling effects that might have been observed at baseline would diminish, as well as to test for b) transfer effects within the kindergarten groups, we conducted univariate ANOVAs to test for group effects at each session, which were followed by post-hoc tests: a) CG vs. SG, WM vs. SG, IC vs. SG; b) WM vs. CG, IC vs. CG, as well as WM vs. IC.

To conduct our analyses, we used JASP (JASP, 2017) and report Bayes Factors (BF) for each of our analyses. We calculated BF relying on default priors as recommended by Marsman & Wagenmakers (2017) and Wagenmakers et al. (2017). Specifically, we report the probability of our data fitting under the null vs. the alternative hypothesis (note that values larger than

---

[2] The two versions did not differ from each other, at least in our sample, i.e. versions A and B resulted in similar scores in the kindergarten participants at baseline ($BF_{10} = 0.29$).

one are in favor of the hypothesis and that values smaller than one are in favor of the null; see also Jarosz & Wiley, 2014).

Results

Descriptive data for all assessment sessions and groups are provided in Table 1.

Baseline Performance – Kindergarten Group Matching and Schooling Effect

Comparing the three kindergarten groups, there were no group differences in most of the measures (BDST: $BF_{10} = 0.42$; BSI RT: $BF_{10} = 0.13$; Gf: $BF_{10} = 0.15$), indicating that the groups were well matched at baseline in those measures. However, there was a substantial group difference for the BSI errors in the AX-CPT ($BF_{10} = 23.87$), and thus, we ran additional post-hoc tests which revealed that the WM group outperformed the two other groups (WM vs. IC: $BF_{10} = 11.16$; WM vs. CG: $BF_{10} = 17.63$), but there was no difference between the two remaining groups (IC vs. CG: $BF_{10} = 0.30$).

Next, comparing the schooling group with all kindergarten groups to investigate the schooling effect, there was no evidence for group differences in the BDST ($BF_{10} = 0.49$) and AX-CPT (BSI RT: $BF_{10} = 0.27$). However, for Gf, there was very strong evidence for a schooling effect in that the schooling group outperformed the kindergarten groups ($BF_{10} = 58.69$; see Figure 3b). For AX-CPT (BSI errors), given that the WM group outperformed the IC and the kindergarten control group at baseline, we calculated two separate ANOVAs, one comparing the schooling group with the WM group, which did not provide any evidence for a schooling effect ($BF_{10} = 0.77$), but comparing the schooling group with the remaining two kindergarten groups revealed anecdotal to moderate evidence for a schooling effect (SG vs. IC and CG: $BF_{10} = 3.07$).

Training Effects

The average performance of all trainees during the five-week intervention period as a function of training task is presented in Figure 2. The children in the WM group decisively improved their performance as reflected by their increase in maximum set sizes ($BF_{-0} = 2.39e+6$), as did the children in the IC training group in both stop-signal task (SSRT: $BF_{+0} = 4.53$), as well as the modified Stroop task (accuracy in the congruent condition: $BF_{-0} = 115.17$; accuracy in the incongruent condition: $BF_{-0} = 14{,}577.89$; RT in the congruent condition: $BF_{+0} = 8.81e+6$; RT in the incongruent condition: $BF_{+0} = 657{,}268.79$).

Transfer Effects

WM Task.—A mixed ANOVA on the BDST to test for immediate changes (pre vs. post) revealed no evidence for either the main effect of session ($BF_{10} = 0.36$), or the main effect of group ($BF_{10} = 0.16$), or the session by group interaction ($BF_{10} = 0.15$). Within-group comparisons did not indicate any changes either (all $BF_{-0} \leq 0.60$), except for the IC group, but the evidence was merely anecdotal ($BF_{-0} = 1.92$) (see Table 1).

A mixed ANOVA to test for any longitudinal effects (BDST; pre vs. follow-up) revealed a decisive main effect of session ($BF_{10} = 564.04$), as well as a moderate main effect of group

($\mathrm{BF}_{10} = 3.88$), however, there was no evidence for an interaction between session and group ($\mathrm{BF}_{10} = 0.28$). Within-group comparisons showed that the schooling group and the IC training group substantially improved from baseline to follow-up (SG: $\mathrm{BF}_{-0} = 7.43$; IC: $\mathrm{BF}_{-0} = 12.29$), while the evidence was merely anecdotal in the WM training group ($\mathrm{BF}_{-0} = 2.70$). The kindergarten control group did not show any change ($\mathrm{BF}_{-0} = 0.37$; see Table 1 and Figure 3a).

Similarly, the ANOVAs that specifically compared the WM or IC training groups with kindergarten control group revealed anecdotal evidence for an interaction between session and group, but only from pretest to follow-up, and only for the IC vs. CG comparison (IC vs. CG: $\mathrm{BF}_{10} = 30$; all other $\mathrm{BF}_{10} < 1$).

To further test for group differences, we compared the performance of the four groups at each test session. There was no evidence for group differences at either the pretest ($\mathrm{BF}_{10} = 0.34$) or the posttest ($\mathrm{BF}_{10} = 0.08$). However, strong evidence for group differences appeared at follow-up ($\mathrm{BF}_{10} = 31$), and further analyses revealed that the schooling group decisively outperformed the kindergarten control group ($\mathrm{BF}_{10} = 122.81$), providing strong evidence for a schooling effect. Notably, at this point also the two training groups outperformed the kindergarten control group indicating some evidence for transfer. While there was strong evidence for an effect in the WM training group (WM vs. CG: $\mathrm{BF}_{10} = 25.77$), the evidence was merely anecdotal for the IC group (IC vs. CG: $\mathrm{BF}_{10} = 1.94$). Overall, those results provide tentative evidence that schooling and WM training impact BDST performance over the long run, and notably, there were no differences between the two training groups and the schooling group (all $\mathrm{BF}_{10} \leq 0.55$), suggesting that both targeted training and schooling impact BDST performance (see Figure 3a).

IC Task.—The mixed ANOVA to test for immediate changes in the AX-CPT (BSI errors; pre vs. post) revealed no evidence for a main effect of session ($\mathrm{BF}_{10} = 0.57$), but there was very strong evidence for a main effect of group ($\mathrm{BF}_{10} = 38.77$), however, there was no evidence for a session by group interaction ($\mathrm{BF}_{10} = 0.16$). Within group comparisons indicated that only the schooling group substantially changed from baseline to posttest ($\mathrm{BF}_{-0} = 7.47$; i.e. changing towards using a more proactive strategy), while none of the three kindergarten groups did (all $\mathrm{BF}_{-0} \leq 0.83$).

The mixed ANOVA to test for the longitudinal effects (BSI errors; pre vs. follow-up) provided decisive evidence for a main effect of session ($\mathrm{BF}_{10} = 407.00$), as well as strong evidence for a main effect of group ($\mathrm{BF}_{10} = 26.11$). However, there was only anecdotal evidence for an interaction between group and session ($\mathrm{BF}_{10} = 1.14$). Within group comparisons indicated that the schooling group as well as the kindergarten control group changed from baseline to follow-up (SG: $\mathrm{BF}_{-0} = 324$; CG: $\mathrm{BF}_{-0} = 12.86$), while there was no evidence for changes in the two trainings groups (both $\mathrm{BF}_{-0} \leq 0.60$).

This pattern was further illustrated by the fact that the ANOVAs that specifically compared the WM or IC training group with kindergarten control group revealed anecdotal evidence for session by group interactions, but only from pretest to follow-up, which were driven by the changes in kindergarten control group (IC vs. CG: BF$_{10}$ = 1.66; WM vs. CG: BF$_{10}$ = 1.29).

Comparing the performance between the four groups revealed group differences at pretest (BF$_{10}$ = 19.22), and further analyses indicated anecdotal evidence for a schooling effect in that the schooling group differed from the kindergarten control group and the IC training group (SG vs. CG: BF$_{10}$ = 1.94; SG vs. IC: BF$_{10}$ = 1.40). In addition, there was strong evidence showing that the WM training group differed from both, the kindergarten control group (WM vs. CG: BF$_{10}$ = 17.63) and IC training group (WM vs. IC: BF$_{10}$ = 11.16). There was no evidence for any other group differences (all BF$_{10}$ $\leq$ 0.77; see Figure 3b).

There was also substantial evidence for group differences at posttest (BF$_{10}$ = 4.90). Further analyses showed a similar pattern that was observed at pretest, indicating anecdotal to moderate evidence for a schooling effect in that the schooling group was still different from the kindergarten control group (SG vs. CG: BF$_{10}$ = 1.30) and the IC training group (SG vs. IC: BF$_{10}$ = 5.30). Furthermore, the WM training group was still different from both, the kindergarten control group (albeit weakly) and the IC training group (WM vs. CG: BF$_{10}$ = 1.72; WM vs. IC: BF$_{10}$ = 5.72). There was no evidence for other group differences (all BF$_{10}$ $\leq$ 0.39).

At the follow-up test, the overall group differences remained substantial (BF$_{10}$ = 7.10). Further analyses revealed strong to decisive evidence that the schooling group and the WM training group were different from the IC training group (SG vs. IC: BF$_{10}$ = 107; WM vs. IC: BF$_{10}$ = 11.52). There was also anecdotal evidence that the kindergarten control group was now also different from the IC training group (CG vs. IC: BF$_{10}$ = 1.34). There was no evidence for any other group differences (all BF$_{10}$ $\leq$ 0.41).

To further explicate the children’s strategy use, we calculated the difference between BSI and 0, which indicates whether the children adopted a more proactive (> 0) or reactive (< 0) strategy (see Figure 3b). At pretest, there was anecdotal evidence that the WM training group was relying on a proactive strategy (BSI > 0; BF$_{10}$ = 1.86), whereas the IC training group and the kindergarten control group relied on a reactive strategy (BSI < 0; IC: BF$_{10}$ = 1.39; CG: BF$_{10}$ = 2.31). The score in the schooling group was not different from 0 (BF$_{10}$ = 0.25). At posttest, there was now moderate evidence that the WM training group and the schooling group relied on a proactive strategy (BSI > 0; WM: BF$_{10}$ = 3.77; SG: BF$_{10}$ = 3.87). The BSI in the IC training group and kindergarten control group was no longer different from 0 (both BF$_{10}$ $\leq$ 0.58). At follow-up, the evidence for the WM training group and the schooling group using a proactive strategy was now even stronger, and it was decisive in the case of the schooling group (BSI > 0; WM: BF$_{10}$ = 12.20; SG: BF$_{10}$ = 3,403). In contrast, the BSI for neither the IC training group nor the kindergarten control group was different from 0 (both BF$_{10}$ $\leq$ 0.67).

The mixed ANOVA to test for immediate changes (pre vs. post) for the BSI RT revealed no evidence for a main effect of session (BF$_{10}$ = 0.17) or for a main effect of group (BF$_{10}$ = 0.35), or for a session by group interaction (BF$_{10}$ = 0.41), and the within group comparisons revealed no changes from baseline to the posttest (all BF$_{-0}$ $\leq$ 0.58).

To test for the longitudinal effects (BSI RT; pre vs. follow-up), the mixed ANOVA indicated anecdotal evidence for a main effect of session ($BF_{10} = 1.35$), but there was no evidence for either the main effect of group ($BF_{10} = 0.06$), or the group by session interaction ($BF_{10} = 0.08$), and the within group comparisons did not indicate any changes from baseline to follow-up (all $BF_{-0} \leq 0.18$).

The ANOVAs that specifically compared the WM or IC training with kindergarten control group revealed anecdotal evidence for a session by group interaction, but only from pretest to posttest, and only for the WM vs. CG comparison (WM vs. CG: $BF_{10} = 1.79$; all other $BF_{10} < 1$).

Comparing the performance of the four groups at each test session provided no evidence for group differences at pretest ($BF_{10} = 0.07$). However, there was anecdotal evidence for group differences at posttest ($BF_{10} = 2.28$). Further analyses showed that both, the WM training group and the IC training group were different from the kindergarten control group (WM vs. CG: $BF_{10} = 4.73$; IC vs. CG: $BF_{10} = 2.88$), and there was anecdotal evidence for a difference between the schooling group and the kindergarten control group as well (SG vs. CG: $BF_{10} = 1.58$). There was no evidence for other group differences (all $BF_{10} \leq 0.55$). At follow-up, there was no evidence for group differences ($BF_{10} = 0.09$) (see Figure 3c).

Gf Task.—With regards to the immediate changes (pre vs. post), the mixed ANOVA revealed anecdotal evidence for a main effect of session ($BF_{10} = 1.77$) and moderate evidence for a main effect of group ($BF_{10} = 5.55$). However, there was no evidence for a session by group interaction ($BF_{10} = 0.12$). Within group comparisons indicated anecdotal evidence for the WM group improving from baseline to posttest ($BF_{-0} = 2.65$), but there was no evidence for any of the other groups improving (all $BF_{-0} \leq 0.68$).

With respect to the longitudinal effect (pre vs. follow-up), the mixed ANOVA revealed decisive evidence for a main effect of session ($BF_{10} = 327.24$) as well as very strong evidence for a main effect of group ($BF_{10} = 65.04$). However, there was no evidence for an interaction between session and group ($BF_{10} = 0.93$). Nonetheless, it is worth noting that both training groups substantially improved from baseline to follow-up (WM: $BF_{-0} = 17.12$; IC: $BF_{-0} = 19.74$), as did the schooling group, although to a lesser extent ($BF_{-0} = 3.15$). In contrast, the kindergarten control group did not show any change ($BF_{-0} = 0.29$; see Table 1). To further illustrate this point, the ANOVAs that directly compared the WM or IC training group with kindergarten control group revealed anecdotal evidence for session by group interactions, but only from pretest to follow-up (WM vs. CG: $BF_{10}= 1.78$; IC vs. CG: $BF_{10} = 1.29$).

Comparing the performance of the four groups in each test session, there was moderate evidence for a group effect at pretest ($BF_{10} = 8.00$). Further analyses indicated that the schooling group outperformed each of the three kindergarten groups (SG vs. WM: $BF_{10} = 6.56$; SG vs. IC: $BF_{10} = 4.74$; SG vs. CG: $BF_{10} = 22.23$), indicating a schooling effect. At posttest, although there was no evidence for an overall group difference ($BF_{10} = 0.46$), further analyses revealed anecdotal evidence for the schooling group still outperforming the kindergarten control group ($BF_{10} = 1.92$). There was no evidence for any other group difference at posttest (all $BF_{10} \leq 0.95$). At follow-up, there was very strong evidence for a group difference ($BF_{10} = 65.55$). Further analyses revealed that the schooling group now decisively outperformed the kindergarten control group (SG vs. CG: $BF_{10} = 663.42$), however, there was now also substantial evidence for the two training groups outperforming the kindergarten control group as well (WM vs. CG: $BF_{10} = 7.01$; IC vs. CG: $BF_{10} = 5.55$), indicating transfer in that both, WM and IC training facilitated performance in the Gf task. Importantly, there was no longer any evidence for a difference between the two training groups and the schooling group that was observed at baseline (both $BF_{10} \leq 0.80$), despite the fact that the training groups now also outperformed the kindergarten control group control group, indicating that the training groups caught up to the schooling group (see Figure 3d).

Discussion

The present study had two main goals: First, to replicate and extend the schooling effects to various measures of EF and Gf, and second, to investigate whether targeted training could lead improvements in EF and Gf, and if so, whether the improvements would be similar to those observed as a result of schooling. Using a school-cutoff design, we trained kindergartners targeted either WM or IC skills for 20 sessions over 5 weeks, and we compared their performance with that of age-matched first-graders receiving their regular school curriculum, as well as with a business-as-usual kindergarten control group. All children were tested on measures of WM, IC, and Gf three times at baseline, posttest and 3 months later for a follow-up. Our results demonstrated a schooling effect, which was especially pronounced in Gf. Furthermore, children in both training groups did not only improve on their respective trained tasks, but also showed varying degrees of improvement in the untrained measures. Most importantly, both training groups ultimately caught up with the schooling group in that their Gf performance was comparable to that of the first graders at follow-up.

Schooling effect

Our results indicated that there are beneficial effects of schooling on specific cognitive functions. In particular, our age-matched first graders outperformed all kindergartners at baseline in Gf. Furthermore, this difference persisted when comparing the first graders and kindergarten control group over time, with the most pronounced difference observed at follow-up. This finding corroborates earlier work demonstrating a positive relationship between schooling and cognitive ability (Brod, et al., 2017; Ceci, 1991, see also Nisbett, 2013; Nisbett et al., 2012).

There were some additional indications for schooling effects, although they were less pronounced in the other measures. Specifically, with respect to WM, while we did not observe any reliable group differences at baseline or posttest, we did observe group differences at follow-up, indicating that the schooling group now outperformed the kindergarten control group. Finally, while we did observe potential schooling effects in the AX-CPT at baseline (BSI errors) in that the schooling children demonstrated a more proactive control strategy than the kindergarten control and the IC training group, the effect was merely anecdotal, and furthermore, it was complicated by the fact that the WM training group showed an even more pronounced proactive strategy at baseline (i.e. showing BSI scores that were larger than 0; Figure 3b). This pattern remained the same at posttest, however, at follow-up, also the kindergarten control group caught up with the schooling and WM training groups, the changes over the course of the study might reflect a general maturation and/or regression to the mean phenomenon rather than schooling (or targeted training), which is consistent with earlier work showing that not all EF seem to be susceptible to the effects of schooling (Burrage et al., 2008). Nonetheless, the shift observed in the kindergarten control group (Figure 3b) is consistent with previous findings reporting a maturational shift from reactive to proactive control during that age (Lucenet & Blaye, 2014; Chevalier et al., 2014).

WM training

Children in the WM training group improved decisively in the trained task, which is in line with previous research (Karbach et al., 2015; Loosli et al., 2012). Although children in the WM training group did not show any reliable improvement in BDST immediately after training, there was anecdotal evidence for an improvement at follow-up 3 months later, and moreover, the WM training group outperformed the control group at this time, however, there was no session x group interaction. Even though this improvement in BDST seems modest given that this group specifically trained on WM, two issues might have contributed to this result: first, the WM training group started with rather high baseline scores, and as such, their ability to improve might have been limited, and second, the BDST requires additional executive control resources that go beyond the more basic WM requirements of the training task (St Clair-Thompson, 2010), which might also explain why the IC group showed improvements in the BDST (see below). However, we did observe improved performance in Gf as a function of WM training as compared to the kindergarten control group, but again, those benefits were only present at follow-up, as evidenced by a substantial group difference between the WM training group and the kindergarten control group at follow-up, which was somewhat mitigated by the anecdotal evidence for an interaction. Importantly, even though the schooling group outperformed the WM training group at baseline, there was no more evidence for such group differences after training.

In contrast, we did not see any changes in AX-CPT performance in the WM training group. Overall, our results suggest that both, targeted WM training and schooling lead to improvements in important cognitive measures, however, those effects seem to require time to manifest themselves given that they were only observed at follow-up.

Even though those longitudinal findings might seem counterintuitive, similar effects have been reported using a variety of interventions and described as ‘sleeper’ effects (e.g., in psychotherapy research, Moritz, et al., 2014; Van Aar, Leijten, Orobio de Castro, & Overbeek, 2017). In the cognitive training field, Van der Molen et al. (2010) found that transfer effects from WM training to short-term memory, scholastic abilities and story recall appeared ten weeks after training in adolescents with mild to borderline intellectual disabilities. In another study targeting children with low WM skills, Holmes, Gathercole, & Dunning (2009) reported improvements in math 6 months later. But the control group in this study was not re-tested, and as such, it is unclear whether the gains were due to training or simply a function of age-related developmental changes. Blair and Raver (2014) reported significant long-term benefits of Tools of the Mind at the end of first grade instead of the end of Kindergarten. Despite the fact that the immediate effects are small or non-existent, one possibility for long-term benefits is that training-induced cognitive improvement might need time to result in measurable behavioral effects, which might be related to consolidation effects (Au, Karsten, Buschkuehl, & Jaeggi, 2017). Indeed, several studies have shown that a testing delay enhances children's learning and memory formation, eliminating interference and indicating consolidation (Darby & Sloutsky, 2015; Ellenbogen, Hu, Payne, Titone, & Walker, 2007; Werchan & Gómez, 2013). Another possibility that does not exclude the previous one could be that children learn how to implement and translate their newly acquired skills in other situations over time, indicating changes in ‘underlying competence’ that go beyond mere practice effects (Klauer & Phye, 2008). Thus, we speculate that children might indeed need time to "digest" instead of absorb the generalized cognitive ability during or even immediately after training. Furthermore, given that there was no significant improvement in the kindergarten control group from pretest to follow-up test in either BDST or Gf, our effects seem to go beyond simple maturation processes, suggesting that the changes observed in the WM training group were most likely due to the training itself.

Although we acknowledge that many WM training studies fail to observe improvements in Gf, including one of our own in which we used the same intervention (Loosli et al., 2012), there are several studies that do show improvements in both children and young adults (Jaeggi et al., 2008; Jaeggi, Buschkuehl, Shah, & Jonides, 2014; Klingberg et al., 2002, 2005). One of the reasons we might not have observed reliable improvements in Gf in our earlier study might be because the training time in that study was considerably shorter, most importantly however, our earlier study did not assess long-term effects, and as discussed above, the transfer effects might not have been apparent at posttest.

IC training

Children in the IC training group decisively improved their performance in the modified Stroop task, and they also showed improvements in the Stop-signal task, although to a lesser extent. Such training-specific improvements are consistent with previous findings that demonstrated improvements in stop-signal task performance (Berkman et al., 2014), go/no-go and flanker performance (Thorell et al., 2009), as well as Stroop performance (Dulaney & Rogers, 1994; MacLeod, 1998; MacLeod & Dunbar, 1988).

Despite those training-specific improvements, there is not much literature suggesting transfer to related tasks as a function of IC training, especially in young children. For example, while training on several attention tasks including a Stroop-like task in preschoolers led to neural changes associated with a flanker task (Rueda et al., 2005, 2012), no behavioral effects were observed, and another IC training program consisting of go/no-go, flanker, and stop-signal tasks did not generalize to a Stroop-like task in preschoolers (Thorell et al., 2009).

As such, it might not be surprising that we found very little evidence for improvements in the AX-CPT BSI score as a result of IC training, in fact, the IC training group was the only group that kept relying on a relatively reactive strategy despite (or maybe because) receiving targeted training. One could speculate that our intervention might have inadvertently discouraged children to use a proactive control strategy, leading to an increase in BX errors and a decrease in AY errors (Table S1), along with relatively lower BSI error scores in this group, and as such, preventing the maturational shift from reactive to proactive processes (Gonthier et al., 2016; Lucenet & Blaye, 2014; Chevalier et al., 2014). That is, while the performance in AX-CPT benefits from a proactive strategy (i.e. response preparation), performance in the trained stop-signal task benefits from a more reactive strategy in that participants have to withhold their response upon seeing an (unpredictable) stop stimulus.

Interestingly, the IC training group showed strong evidence for performance improvements in both WM and Gf, and they further outperformed the kindergarten control group, but as with the WM training group, this effect became only apparent at follow-up. Furthermore, despite that the schooling group outperformed the IC training group at baseline in Gf, there was no longer any evidence for that difference at follow-up, suggesting again that both groups benefitted equally from their experiences.

Even though it has been difficult to show effects that go beyond specific effects after IC training (Thorell et al., 2009), there have been occasional reports of transfer, for example, Rueda et al. (2012) trained 5-year old children on a set of executive attention games that included Stroop-like tasks, and they observed a significant gain in the attention network test (ANT) as well as in measures on Gf. Also, a recent study by Zhao et al. (2018) with older children (10-12-year-olds) observed improvements in WM after go/no-go inhibition training, which was most pronounced after 3 months, which is consistent with our own finding, emphasizing the need for researchers to routinely include follow-up assessments, especially in children populations.

Although Zhao and colleagues (2018) argued that their (albeit modest) success might have been due to the fact that they relied on only one training task instead of a battery of tasks, it has generally been argued that training variability is critical for learning outcome (Schmidt & Bjork, 1992). In our case, although the overall training time of two inhibition control tasks amounted to the same training time of our WM training task, our IC training group experienced more training variability due to the fact that they trained on two different tasks each day. Whether or not this variable approach was more successful than the 1-task approach we used in the WM training group is hard to determine with our data.

Limitations

We observed no evidence for changes in the BSI RT measure (AX-CPT) in any of the groups, and it is important to note that the re-test reliabilities were fairly low (see Table 1). Although issues with reliability are not uncommon when assessing cognitive functions in early childhood (e.g., Ramani et al., 2017), the fact that we used difference scores consisting of measures that were made up by relatively few trials might have been especially detrimental for reliability (see also Table S1).

Another caveat that might have contributed to the relatively small effects is that our population consisted of typically developing children from predominantly privileged backgrounds, and thus, our results might have been affected by restricted range issues. There is evidence that participants with lower baseline ability levels typically improve more in training and transfer than their high-ability peers (e.g., Jaeggi et al., 2011; Jaeggi, Karbach, & Strobach, 2017; Karbach, Koenen, & Spengler, 2017), and furthermore, children from low SES backgrounds seem to show especially large improvements in EF after targeted training, presumably since they are at a higher risk of having deficits in those areas, and consequentially, they have more room to improve (Blair & Raver, 2014; Segretin et al., 2014). Due to time restrictions with testing, it is important to note that we only implemented one test per cognitive domain and thus, we are not able to determine whether any improvements would hold on a latent level, that is, whether our results would extend beyond those specific tasks that we used here.

Finally, we acknowledge that our study might have benefitted from a larger sample size and from including at least an additional, active, control group, however, we were constrained by the number of students in that particular school district that fit our exact age criterion. Overall, future studies with larger samples that include students with wider ability ranges and a broader set of outcome measures would bolster the generalizability of our results.

Overall conclusion and implications

To conclude, our work provides further evidence for the malleability of EF, illustrating the role of specific experiences in driving this malleability. Specifically, our results demonstrate that schooling has beneficial effects on the development of higher cognitive functions. Furthermore, the kindergarten groups that received targeted training did not only show improvements in the trained task, but importantly, they were ultimately able to catch up to the schooling group as a result of their training, which was especially apparent in Gf at follow-up, suggesting that not only long-term experiences such as schooling improve EF in early childhood, but that similar effects can be obtained with short-term cognitive interventions. Overall, our work provides a promising rationale for future work aiming at testing whether targeted EF training might benefit children who lag behind their peers in their regular school environment, and as such, offer a means to improve their readiness to learn.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Q. Z. is supported by the National Natural Science Foundation of China (No. 31100753), National Key Technologies R&D Program (No.2012BAI36B02), the Natural Sciences Foundation of Zhejiang Province (No. LY16C090002), the Social Sciences Foundation of Zhejiang Province (No. 11ZJQN033YB) and the Fundamental Research Funds for the Central Universities (No. 2015QNA3021). S.M.J. is supported by the U.S. National Institute on Aging (No. 1K02AG054665).

We would like to thank the children for their participation, they are from Zhaohui kindergarten, the affiliated Xixi Kindergarten of Zhejiang University and Tianchang Elementary School, Hangzhou, China.

References

Alderson RM, Rapport MD, & Kofler MJ (2007). Attention-deficit/ hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. Journal of Abnormal Child Psychology, 35(5), 745–758. [PubMed: 17668315]

Alloway TP, & Alloway RG (2010). Investigating the predictive roles of working memory and IQ in academic attainment. Journal of Experimental Child Psychology, 106(1), 20–29. [PubMed: 20018296]

Alloway TP, Gathercole SE, Adams AM, Willis C, Eaglen R, & Lamont E (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. British Journal of Developmental Psychology, 23(3), 417–426.

Alloway TP, Gathercole SE, Kirkwood H, & Elliott J (2009). The cognitive and behavioral characteristics of children with low working memory. Child Development, 80(2), 606–621. [PubMed: 19467014]

Au J, Buschkuehl M, Duncan GJ, & Jaeggi SM (2016). There is no convincing evidence that working memory training is NOT effective: A reply to Melby-Lervåg and Hulme (2015). Psychonomic Bulletin & Review, 23(1), 331–337. [PubMed: 26518308]

Au J, Karsten C, Buschkuehl M, & Jaeggi SM (2017). Optimizing Transcranial Direct Current Stimulation Protocols to Promote Long-Term Learning. Journal of Cognitive Enhancement, 1(1), 65–72.

Au J, Sheehan E, Tsai N, Duncan GJ, Buschkuehl M, & Jaeggi SM (2015). Improving fluid intelligence with training on working memory: a meta-analysis. Psychonomic Bulletin & Review, 22(2), 366–377. [PubMed: 25102926]

Baker DP, Eslinger PJ, Benavides M, Peters E, Dieckmann NF, & Leon J (2015). The cognitive impact of the education revolution: A possible cause of the Flynn effect on population IQ. Intelligence, 49,144–158.

Berkman ET, Kahn LE, & Merchant JS (2014). Training-induced changes in inhibitory control network activity. The Journal of Neuroscience, 34(1), 149–157. [PubMed: 24381276]

Blair C, & Raver CC (2014). Closing the achievement gap through modification of neurocognitive and neuroendocrine function: Results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. Public Library of Science One, 9(11), e112393. [PubMed: 25389751]

Blair C, & Razza RP (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child Development, 78(2), 647–663. [PubMed: 17381795]

Braver TS (2012). The variable nature of cognitive control: a dual mechanisms framework. Trends in Cognitive Sciences, 16(2), 106–113. [PubMed: 22245618]

Braver TS, Paxton JL, Locke HS, & Barch DM (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. Proceedings of the National Academy of Sciences of the United States of America, 106(18), 7351–7356. [PubMed: 19380750]

Braver TS, Satpute AB, Rush BK, Racine CA, & Barch DM (2005). Context processing and context maintenance in healthy aging and early stage dementia of the Alzheimer's type. Psychology & Aging, 20(1), 33–46. [PubMed: 15769212]

Brod G, Bunge SA, & Shing YL (2017). Does one year of schooling improve children’s cognitive control and alter associated brain activation? Psychological Science, 28(7), 967–978. [PubMed: 28489500]

Burrage MS, Ponitz CC, McCready EA, Shah P, Sims BC, Jewkes AM, & Morrison FJ (2008). Age- and schooling-related effects on executive functions in young children: A natural experiment. Child Neuropsychology, 14(6), 510–524. [PubMed: 18982508]

Carlson SM, Zelazo PD, & Faja S (2013). Executive function. In Zelazo PD (Ed.), The Oxford Handbook of Developmental Psychology: Vol. 1, Body and mind (pp. 706–743). New York: Oxford University Press.

Ceci S (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. Developmental Psychology, 27, 703–722.

Chein JM, & Morrison AB (2010). Expanding the mind’s workspace: Training and transfer effects with a complex working memory span task. Psychonomic Bulletin & Review, 17(2), 193–199. [PubMed: 20382919]

Chevalier N, James TD, Wiebe SA, Nelson JM, & Espy KA (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), 1999–2008. [PubMed: 24773104]

Chiew KS, & Braver TS (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. Cognitive, Affective, & Behavioral Neuroscience, 14(2), 509–529.

Christoff K, Ream JM, Geddes L, & Gabrieli JD (2003). Evaluating self-generated information: anterior prefrontal contributions to human cognition. Behavioral Neuroscience, 117(6), 1161–1168. [PubMed: 14674837]

Cohen J (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Darby KP, & Sloutsky VM (2015). When Delays Improve Memory Stabilizing Memory in Children May Require Time. Psychological Science, 26(12), 1937–1946. [PubMed: 26525075]

Davidson MC, Amso D, Anderson LC, & Diamond A (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. Neuropsychologia, 44(11), 2037–2078. [PubMed: 16580701]

Diamond A (2013). Executive functions. Annual Review of Psychology, 64, 135–168.

Diamond A, Barnett WS, Thomas J, & Munro S (2007). Preschool program improves cognitive control. Science, 318(5855), 1387–1388. [PubMed: 18048670]

Dulaney CL, & Rogers WA (1994). Mechanisms underlying reduction in Stroop interference with practice for young and old adults. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(2), 470–484

Ellenbogen JM, Hu PT, Payne JD, Titone D, & Walker MP (2007). Human relational memory requires time and sleep. Proceedings of the National Academy of Sciences, 104(18), 7723–7728.

Gathercole SE, Lamont E, & Alloway TP (2006). Working memory in the classroom In Pickering S (Ed.), Working Memory and Education (pp. 219–240). Amsterdam: Elsevier Press.

Gonthier C, Macnamara BN, Chow M, Conway ARA, & Braver TS (2016). Inducing proactive control shifts in the ax-cpt. Frontiers in Psychology, 7: 1822 [PubMed: 27920741]

Holmes J, Gathercole SE, & Dunning DL (2009). Adaptive training leads to sustained enhancement of poor working memory in children. Developmental Science, 12(4), F9–F15. [PubMed: 19635074]

Jaeggi SM, Buschkuehl M, Jonides J, & Perrig WJ (2008). Improving fluid intelligence with training on working memory. Proceedings of the National Academy of Sciences, 105(19), 6829–6833.

Jaeggi SM, Buschkuehl M, Jonides J, & Shah P (2011). Short-and long-term benefits of cognitive training. Proceedings of the National Academy of Sciences, 108(25), 10081–10086.

Jaeggi SM, Karbach J, & Strobach T (2017). Editorial Special Topic: Enhancing brain and cognition through cognitive training. Journal of Cognitive Enhancement, 1(4), 353–357. [PubMed: 29552678]

Jarosz A, & Wiley J (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. The Journal of Problem Solving, 7(1): 2.

Karbach J, & Verhaeghen P (2014). Making Working Memory Work: A Meta-Analysis of Executive-Control and Working Memory Training in Older Adults. Psychological Science, 25(11), 2027–2037. [PubMed: 25298292]

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. [PubMed: 24697256]

Klauer KJ, & Phye GD (2008). Inductive reasoning: A training approach. Review of Educational Research, 78(1), 85–123.

Klingberg T, Fernell E, Olesen PJ, Johnson M, Gustafsson P, Dahlström K, Gillberg CG, 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. [PubMed: 15689731]

Klingberg T, Forssberg H, & Westerberg H (2002). Training of working memory in children with ADHD. Journal of Clinical and Experimental Neuropsychology, 24(6), 781–791. [PubMed: 12424652]

Kroesbergen EH, Van't Noordende JE, & Kolkman ME (2012). Number sense in low-performing kindergarten children: Effects of a working memory and an early math training Reading, Writing, Mathematics and the Developing Brain: Listening to Many Voices. Springer Netherlands, 6, 295–313.

Loosli SV, Buschkuehl M, Perrig WJ, & Jaeggi SM (2012). Working memory training improves reading processes in typically developing children. Child Neuropsychology, 18(1), 62–78. [PubMed: 21623483]

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(4): 831. [PubMed: 25120523]

Macleod CM (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. Memory & Cognition, 26(2), 201–211. [PubMed: 9584429]

MacLeod CM, & Dunbar K (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(1), 126–135.

Maraver MJ, Bajo MT, & Gomezariza CJ (2016). Training on working memory and inhibitory control in young adults. Frontiers in Human Neuroscience, 10: 588. [PubMed: 27917117]

Marsman M, & Wagenmakers EJ (2017). Bayesian benefits with JASP. European Journal of Developmental Psychology, 14(5), 545–555

Melby-Lervåg M, & Hulme C (2013). Is working memory training effective? A meta-analytic review. Developmental Psychology, 49(2), 270–291. [PubMed: 22612437]

Melby-Lervåg M, Redick TS, & 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. [PubMed: 27474138]

Minear M, & Shah P (2006). Sources of Working Memory Deficits in Children and Possibilities for Remediation In Pickering S (Ed.), Working Memory and Education (pp. 273–307). Amsterdam: Elsevier Press.

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive Psychology, 41(1), 49–100. [PubMed: 10945922]

Moritz S, Veckenstedt R, Andreou C, Bohn F, Hottenrott B, & Leighton L, et al. (2014). Sustained and "sleeper" effects of group metacognitive training for schizophrenia: a randomized clinical trial. Jama Psychiatry, 71(10), 1103–1111. [PubMed: 25103718]

Morrison FJ, Smith L, & Dow-Ehrensberger M (1995). Education and cognitive development: A natural experiment. Developmental Psychology, 31(5), 789–799.

Nisbett RE (2013). Schooling makes you smarter: What teachers need to know about IQ. American Educator, 37(1), 10–19.

Nisbett RE, Aronson J, Blair C, Dickens W, Flynn J, Halpern DF, & Turkheimer E (2012). Intelligence new findings and theoretical developments. American Psychologist, 67(2), 130–159. [PubMed: 22233090]

Paxton JL, Barch DM, Storandt M, & Braver TS (2006). Effects of environmental support and strategy training on older adults' use of context. Psychology & Aging, 21(3), 499–509. [PubMed: 16953712]

Prencipe A, Kesek A, Cohen J, Lamm C, Lewis MD, & Zelazo PD (2011). Development of hot and cool executive function during the transition to adolescence. Journal of Experimental Child Psychology, 108(3), 621–637. [PubMed: 21044790]

Ramani GB, Jaeggi SM, Daubert EN, & Buschkuehl M (2017). Domain-specific and domain-general training to improve kindergarten children's mathematics. Journal of Numerical Cognition, 3(2), 468–495.

Raven JC, Court JH, & Raven J (1998). Manual for Raven's Standard Progressive Matrices. Oxford, England: Oxford Psychologists Press.

Raver CC, Jones SM, Li-Grining CP, Metzger M, Champion KM, & Sardin L (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. Early Childhood Research Quarterly, 23(1), 10–26.

Raver CC, Jones SM, Li - Grining C, Zhai F, Bub K, & Pressler E (2011). CSRP's impact on low-income preschoolers' preacademic skills: self - regulation as a mediating mechanism. Child Development, 82(1), 362–378. [PubMed: 21291447]

Rueda MR, Checa P, & Combita LM (2012). Enhanced efficiency of the executive attention network after training in preschool children: immediate changes and effects after two months. Developmental Cognitive Neuroscience, 2, S192–S204. [PubMed: 22682908]

Rueda MR, Rothbart MK, McCandliss BD, Saccomanno L, & Posner MI (2005). Training, maturation, and genetic influences on the development of executive attention. Proceedings of the national Academy of Sciences of the United States of America, 102(41), 14931–14936. [PubMed: 16192352]

Schmidt RA, & Bjork RA (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychological Science, 3(4), 207–217.

Schmiedek F, Lövdén M, & Lindenberger U (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. Frontiers in Aging Neuroscience, 2: 27. [PubMed: 20725526]

Segretin MS, Lipina SJ, Hermida MJ, Sheffield TD, Nelson JM, Espy KA, & Colombo JA (2014). Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds. Frontiers in Psychology, 5: 205. [PubMed: 24659975]

Soveri A, Antfolk J, Karlsson L, Salo B, & Laine M (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. Psychonomic Bulletin & Review, 1–20. [PubMed: 27368622]

St Clair-Thompson Helen L. (2010). Backwards digit recall: A measure of short-term memory or working memory? European Journal of Cognitive Psychology, 22: 2: 286–296.

Thorell LB, 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. [PubMed: 19120418]

Van Aar, Leijten P, Orobio de Castro B, Overbeek G (2017). Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. Clinincal Psychology Review, 51, 153–163

Van der Molen M, Van Luit JEH, Van der Molen MW, Klugkist I, & Jongmans MJ (2010). Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. Journal of Intellectual Disability Research, 54(5), 433–447. [PubMed: 20537049]

Verbruggen F, Chambers CD, & Logan GD (2013). Fictitious inhibitory differences: how skewness and slowing distort the estimation of stopping latencies. Psychological Science, 24(3), 352–362. [PubMed: 23399493]

Wagenmakers EJ, Love J, Marsman M, Jamil T, Ly A, Verhagen J, et al., (2017). Bayesian inference for psychology. Part II: Example applications with JASP. Psychonomic Bulletin & Review, 1–19. [PubMed: 27368622]

Wass SV, Scerif G, & Johnson MH (2012). Training attentional control and working memory–Is younger, better? Developmental Review, 32(4), 360–387.

Wechsler D (1991). WISC-III: Wechsler intelligence scale for children: Manual. Psychological Corporation.

Weicker J, Villringer A, & Thone-Otto A (2016). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injuredpatients. Neuropsychology, 30(2), 190–212. [PubMed: 26237626]

Werchan DM, & Gómez RL (2013). Generalizing memories over time: Sleep and reinforcement facilitate transitive inference. Neurobiology of Learning and Memory, 100, 70–76. [PubMed: 23257278]

Wexler BE, Iseli M, Leon S, Zaggle W, Rush C, & Goodman A, et al. (2016). Cognitive priming and cognitive training: immediate and far transfer to academic skills in children. Scientific Reports, 6, 32859. [PubMed: 27615029]

Willoughby MT, Kupersmidt JB, & Voegler-Lee ME (2012). Is preschool executive function causally related to academic achievement? Child Neuropsychology, 18(1), 79–91. [PubMed: 21707258]

Zelazo PD, & Carlson SM (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. Child Development Perspectives, 6(4), 354–360.

Zhao X, Chen L, & Maes JHR (2018). Training and transfer effects of response inhibition training in children and adults. Developmental Science, 21(1), 1–12.

Research highlights

- We observed schooling effects which were most pronounced in a test of fluid reasoning by comparing first graders and kindergartners with similar chronological age.
- Targeted training that focused on either working memory or inhibitory control led to improvements in a test of fluid reasoning as observed three months after training completion.
- Kindergartners who received targeted training were ultimately able to catch up with their first-grade peers with regards to their fluid reasoning performance.

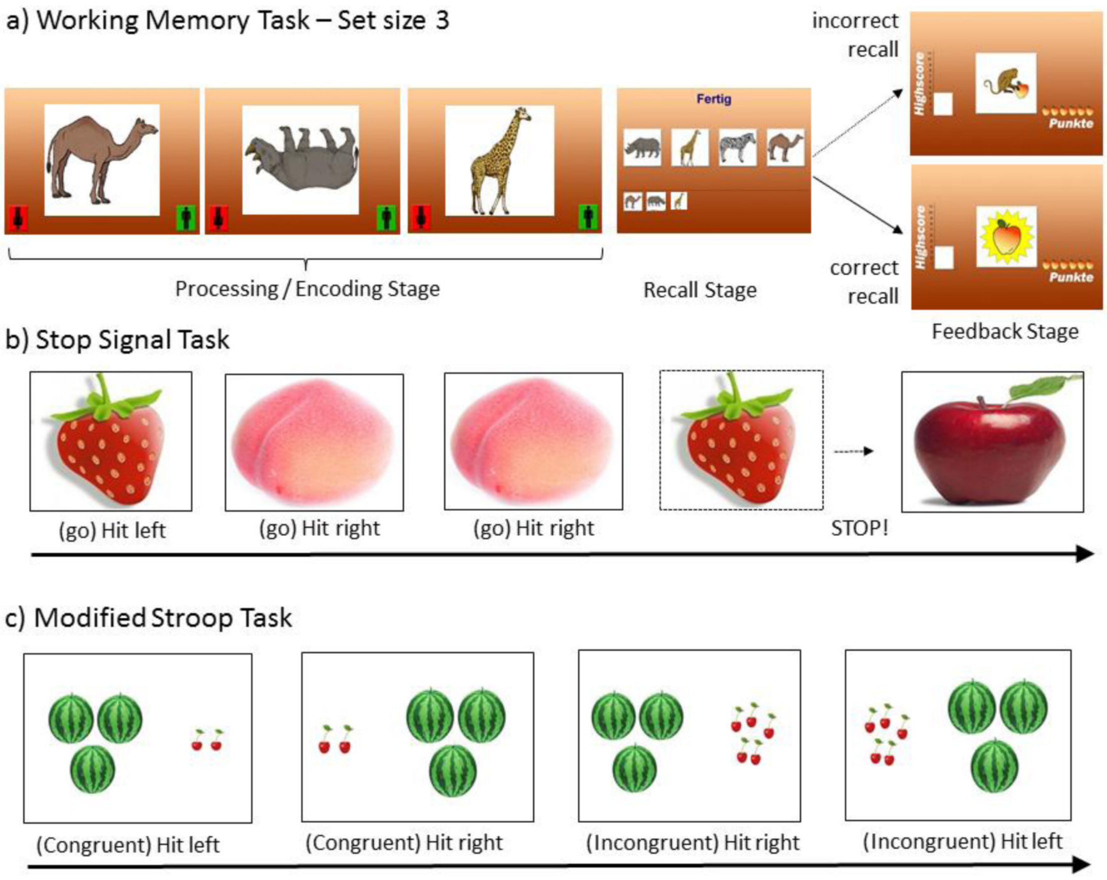

Figure 1.

Training Tasks (see text for task description). a) An example of the working memory training task with set size 3; b) Example trials for the Stop-signal task; c) Example trials for the modified Stroop task.

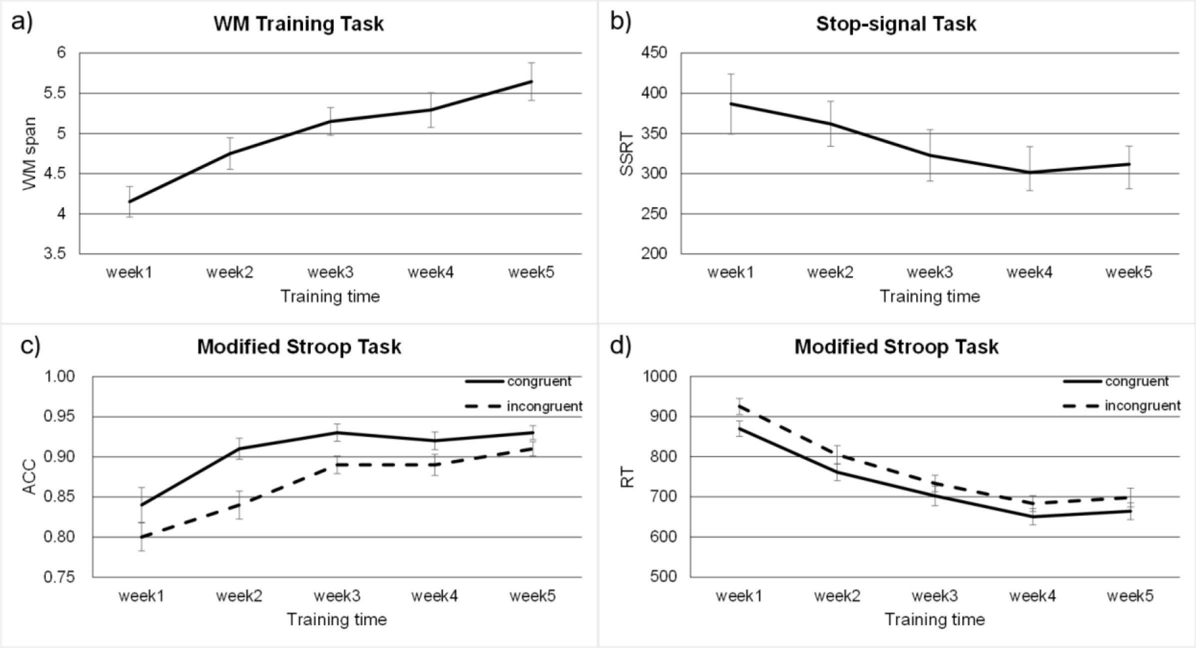

Figure 2.

Training performance in the trained tasks; a) WM span for the WM Training Task, b) SSRT for the Stop-signal task, c) Accuracy for the modified Stroop task, d) Reaction time for the modified Stroop task; error bars refer to standard errors of the mean.

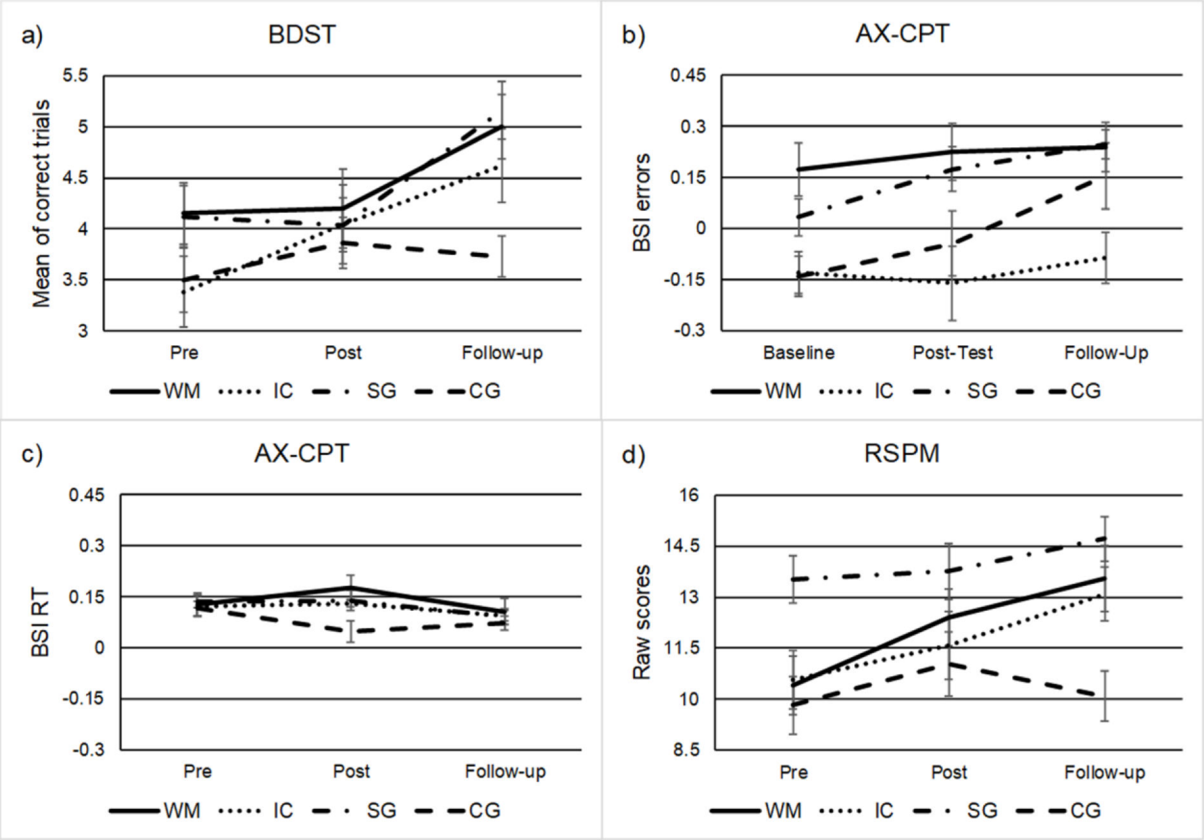

Figure 3.

Task performance across the three test sessions and as a function of group: a) The average correct trials in Backward Digit Span Task; b) The BSI errors in AX-CPT; c) The BSI RT in the AX-CPT. d) The correct responses in the Raven Standard Progressive Matrices Test.

Table 1.

Descriptive data and effect sizes, as well as results for paired comparisons within each group.

	Descriptive Data							Paired Comparisons					
		Pre		Post		Follow-Up		Pre vs. Post			Pre vs. Follow-up		
	N	Mean	SD	Mean	SD	Mean	SD	BF_{-0}	r	ES	BF_{-0}	r	ES
WM													
BDST (correct trials)													
WM	20	4.15	1.35	4.20	1.74	5.00	1.41	0.25	0.26	0.03	***2.70***	0.14	0.47
IC	21	3.38	1.60	4.05	1.77	4.62	1.66	***1.92***	0.54	0.41	***12.29***	0.30	0.64
SG	25	4.12	1.51	4.04	1.34	5.16	1.40	0.18	0.16	−0.04	***7.43***	0.11	0.54
CG	22	3.50	1.47	3.86	1.16	3.73	0.94	0.60	0.24	0.22	0.37	−0.07	0.13
IC													
AX-CPT (BSI errors)													
WM	20	0.17	0.34	0.23	0.37	0.24	0.32	0.39	0.45	0.14	0.60	0.64	0.23
IC	21	−0.13	0.28	−0.16	0.50	−0.09	0.34	0.19	0.20	−0.06	0.38	0.45	0.13
SG	25	0.03	0.27	0.17	0.32	0.25	0.21	***7.47***	0.62	0.54	***324.22***	0.53	0.89
CG	22	−0.14	0.28	−0.04	0.45	0.15	0.45	0.83	0.63	0.28	***12.86***	0.26	0.64
AX-CPT (BSI RT)													
WM	20	0.13	0.15	0.18	0.16	0.11	0.17	0.58	−0.07	0.22	0.18	0.02	−0.09
IC	21	0.12	0.14	0.13	0.05	0.09	0.07	0.28	−0.08	0.05	0.12	0.44	−0.22
SG	25	0.14	0.10	0.14	0.15	0.09	0.12	0.22	−0.02	0.01	0.09	−0.06	−0.29
CG	22	0.12	0.10	0.05	0.15	0.07	0.10	0.09	0.13	−0.40	0.11	−0.23	−0.27
Gf													
RSPM (correct trials)													
WM	20	10.40	3.86	12.40	3.76	13.55	4.29	***2.65***	0.37	0.47	***17.12***	0.40	0.70
IC	21	10.57	3.97	11.57	4.57	13.10	3.65	0.68	0.56	0.25	***19.74***	0.55	0.70
SG	25	13.52	3.50	13.76	4.09	14.72	3.31	0.27	0.48	0.06	***3.15***	0.68	0.44
CG	22	9.82	4.01	11.05	4.43	10.09	3.49	0.65	0.24	0.24	0.29	0.41	0.07

Note: WM: working memory training group; IC: inhibitory control training group; SG: schooling group; CG: control group; BSI: Behavioral Shift Index. BF_{-0} = Bayes Factor for paired comparisons (note that values >1 that provide evidence in favor of the hypothesis are indicated in bold font); r = correlation between the pre- and posttest/follow-up measures; ES = effect size accounting for r: $(\mu_2 – \mu_1) / \sqrt{(\sigma_1^2 + \sigma_2^2 – 2r_{12}\sigma_1\sigma_2)}$.