A classroom intervention to improve executive functions in late elementary school children: Too "old" for improvements?

Valentin Benzing[1][*], Mirko Schmidt[1], Katja Jäger[2], Fabienne Egger[1], Achim Conzelmann[1] and Claudia M Roebers[2]

[1] Institute of Sport Science, University of Bern, Bremgartenstrasse 145, 3012 Bern, Switzerland
[2] Department of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland

Acknowledgements

Abstract

Background: Given the strong relationship between executive functions and academic achievement, there has been great interest in improving executive functions. School-based group interventions targeting executive functions revealed encouraging results in pre-schoolers and young school children, however, there is a paucity of studies in older primary school children (age 10-12). This is surprising considering that deficits in executive function performance can often be observed in this age group.

Aims: Therefore, the aim of the present study was to examine whether an innovative customized school-based group intervention could improve core executive functions (updating, inhibition, shifting) in the age group concerned.

Sample: In total 118 ten to twelve year old school children were recruited from eight participating classes.

Methods: They were randomly assigned to one of two six week conditions of either a *cognitive games group* comprising of card and board games training executive functions (experimental group), or a *wait list control group* (regular school lessons). In the cognitive games group, the class teachers held a 30-minute training session twice a week. ANCOVAs (using pre-test values as covariate) were used to compare executive function performance between groups.

Results: Results revealed that the cognitive games group improved specific executive functions (updating and shifting) compared to the control group. These findings indicate that a school-based group intervention can improve executive functions, even in "older" primary school children.

Conclusion: The present study provides empirical evidence for the effectiveness of a class-room based cognitive training in older primary school children, and is of practical

relevance for educators.

Introduction

A wide range of literature has demonstrated that executive functions (EF) and academic achievement are closely intertwined throughout childhood development (Best, Miller, & Naglieri, 2011), whereby EF have been shown to predict academic achievement. Hence, children with poor EF are more likely to suffer from a variety of negative outcomes (e.g., behavioural problems; Espy, Sheffield, Wiebe, Clark & Moehr, 2011), including poor academic performance and an increased risk of dropping out of school (Diamond, 2016). Based on this relation, there has been great interest in how to best improve EF (Diamond & Ling, 2016).

The umbrella term EF refers to a set of heterogeneous, higher-order cognitive functions. EF enable controlled, goal-directed and adaptive behaviour. They are utilised in complex and novel situations when paying attention is required (Diamond, 2013). There is a general consensus on what are considered to be the three core EF (Miyake et al., 2000): (1) Updating, defined as the ability to keep information in mind and to mentally process the information (e.g., updating one´s thinking), (2) Inhibition, usually categorized into response inhibition (e.g., suppressing a predominant response) and interference control (e.g., controlling one's attention), and (3) Shifting, including switching between tasks or mental sets (e.g., adjusting to changes in complex task demands). They also build the foundations for higher order EF, such as planning, reasoning or problem solving. These core EF are highly relevant in school settings; for example, when remembering the order of work, or inhibiting the temptation to speak out directly (Diamond, 2013).

EF are critical for academic achievement (Best et al., 2011; Diamond, 2013). A variety of studies have supported this relationship, explaining between 20 and 60 % of the variance in performance (Roebers, 2017). Consequently, EF in early childhood directly predicts later academic achievement (Sjöwall, Bohlin, Rydell, & Thorell, 2017; Viterbori, Usai, Traverso, & De Franchis, 2015), and indirectly influences attention, learning-related behaviour and previously acquired academic skills (Jacobson, Williford, & Pianta, 2011; Neuenschwander, Röthlisberger, Cimeli, & Roebers, 2012). Furthermore, deficits in EF seem to be relatively persistent – can even increase over time – and are associated with a range of negative school-related consequences such as behavioural problems, less self-investment into school and consequently higher drop-out rates (Blair & Diamond, 2008; Espy et al., 2011; Friedman et al., 2007; Vitaro, Brendgen, Larose, & Trembaly, 2005). Given the relative significance of EF, the important question is how best to improve them.

There are many interventions enhancing EF in children and adolescents (for review see Diamond & Ling, in press). These include computer-based training, games, physical activities and school curricula activities (Diamond et al., 2015; Karbach & Unger, 2014; Mackey, Hill, Stone, & Bunge, 2011). There is convincing evidence that EF can be enhanced by these interventions (Diamond & Ling, 2016), however, there are certain characteristics which seem to be crucial to elicit improvements. One is, for example, that the EF must be challenged continuously, calling for an increasing level of difficulty. Another example is that EF improvements are dependent on how an activity is presented and conducted (Diamond & Ling, 2016), where it should be engaging, challenging, and playful.

Manifold school-based interventions are available (for review see Otero, Barker, & Naglieri, 2014), which can be classified into direct or indirect trainings. Direct

trainings which predominantly rely on practice typically target single domains , such as computerized working memory training which specifically stimulates working memory. This approach is associated with greater training improvements within that domain, yet fewer transfer effects are observed. In contrast, indirect trainings, such as strategy learning, mostly target multiple domains. This leads to attenuated training gains, yet substantial transfer effects (Diamond & Ling, in press). Considering this, an intervention that directly taps EF whilst practicing diverse skills in school might promote broad benefits.

School-based interventions in group settings, such as playing games with peers, can combine several advantages from both direct and indirect trainings. Besides its substantial ecological validity, group settings increase EF training opportunities through social interactions (Diamond & Ling, 2016). Additionally, games can be adapted easily be designed to be adaptive, constantly challenging and training specific domains (in this case, the core EF). In other words, continuously changing game situations triggers inhibition of prepotent responses, ensures updating of information, promotes interaction with others, and encourages adaptation to new situations (Blair & Diamond, 2008; Diamond, 2010). Participation in games which involve negotiation skills and the execution of a set of rules is one of the major leisure time activities of primary school children, also known to promote school and social adjustment (Pellegrini, Blatchford, Kato & Baines, 2004).

Previous research has mainly focused on interventions targeting EF for preschool and early primary school children. Röthlisberger, Neuenschwander, Cimeli, Michel, and Roebers (2012), for example, showed that an intervention carried out in kindergarten is feasible and effective in improving EF. However, despite the lack of interventions for the age group, EF have been shown to be crucial for development and

academic success in older primary school children (Best et al., 2011; Latzman, Elkovitch, Young, & Clark, 2010). In fact, there are schoolchildren with considerable deficits in EF performance at any age (Zelazo & Müller, 2010), thus interventions for older primary school children are also needed.

The current study investigated the efficacy of a gamified school-based cognitive group intervention in promoting EF in 10 to 12-year-old primary school children. It was hypothesised that there would be significant increases in the performance of updating, inhibition and shifting in the experimental condition when compared to a wait list control group (regular school lessons).

Methods

Design

In a parallel study design, eight classes were randomly and blindly assigned to one of two six-week conditions by drawing lots. The two conditions were carried out during regular school and consisted of (a) a *cognitive games group;* comprising a gamified EF intervention, or (b) a *wait list control group;* consisting of regular school lessons. After the post tests were finished, the wait list control group received the training material and the training schedule to be able to realize the training at a later time point. The routine class teachers carried out the training. They were blind with regard to the study hypothesis, but informed of the basic aims of the study. EF performance was compared between groups before and after the interventions, by means of computerized assessments.

Participants

A total of 118 children between the ages of 10 to 12 ($M = 11.39$; $SD = 0.68$) from eight different classes participated in the study. The Institutional Review Board of the Faculty

of Human Sciences approved the study protocol. The schools' principals as well as the legal guardians of all subjects provided informed written consent, and the children agreed to participate. Children were informed that they could discontinue at any time during the study. There were some participants with incomplete values at both pre-test (5.9%) and post-test (3.5%) due to illness or incomplete questionnaires. Since Little´s missing completely at random test was not significant (χ^2 (22) = 18.78, p = .659), the resulting missing values were imputed with the help of the expectation–maximization (EM) algorithm, which did not affect the pattern of results. Considering previous studies, a priori power analysis (with 1 – beta error probability = .80; alpha error probability = .05; effect size f = .10; number of groups = 2, number of measurements = 2; and correlation between the repeated measures r = .70) was performed. This resulted in a sample size of 118 subjects. The sample characteristics did not differ between groups (see Table 1).

General Procedure

Preceding the pre-interventional assessment, data on background variables including age, gender, height, and weight, were gathered. In addition, the pubertal status was measured using the pubertal development scale (Watzlawik, 2009). For each gender, it consists of three questions (e.g. for boys, "Have you noticed a deepening of your voice?"). Responses had to be given on a 4-point Likert scale, scoring 1-4 points (e.g., not yet started; barely started; definitely started; seems complete).

As for the cognitive tests, pre- and post-interventional assessments were conducted at the same time of the day (between 10.00 a.m. and 12.00 p.m.). The tests were completed in a quiet room, in groups of four children, supervised by two blinded examiners. The examiners gave general instructions about the cognitive tests, which were then completed on a computer. Children were seated apart from each other to avoid

distractions. They received instructions both on a computer screen and via headphones, which also served as sound absorbers. The examiners were present during the testing to ensure that instructions were followed, and to answer questions.

After the pre-interventional assessment, teachers for the cognitive games condition completed a half-day training program to become familiarised with the principles and rules of the games. Subsequently, the intervention was conducted over a six-week period. To account for implementation accuracy, teachers had to report the number of sessions that had been conducted. According to the teachers, all the planned sessions were carried out.

The training program

Overall, the cognitive intervention lasted six weeks. Every week, two 30-minute sessions were conducted, and the teachers could choose on which two week days the sessions were carried out. During each session, they played two different board or card games in small groups, ensuring each task of the intervention was trained once a week. Thus, there were four different games to play, with three incremental levels of difficulty each. Games began at the lowest level of difficulty which was subsequently increased every two weeks, however, teachers were instructed to flexibly adapt the level of difficulty if children seemed undertaxed.

The cognitive games were derived from previous studies, (Röthlisberger et al., 2012) and specifically adapted to this age group (upon request, the senior author can give a detailed description of the cognitive games). They were based on experimental tasks used in previous literature that quantify the individual differences in EF. Considering the task-impurity problem indicating that EF interact with each other, each game was intended to tap multiple EF. To increase motivation, the four tasks were set in a stimulating context where two children experienced adventures in the rainforest

(Roebers, Röthlisberger, Neuenschwander & Cimeli, 2014). Dependant on the games, they were played in groups of two (medicine woman, cheeky monkey) or four (invasion of insects, fruit salad).

The game "*invasion of insects*" taps interference control and shifting (Manly et al., 2001). It is played with cards depicting different quantities of different types of insects. The cards with the depicted insects must be shuffled and distributed equally to children in the group (face down); one card depicting a glass of marmalade is placed in the middle of the table. To start, children count down from three and must turn a card around. They place the card in front of their stack of cards so that everyone can see their card. Children then look at each other's cards, and sum up the insects of the same kind. Whenever exactly five insects of the same kind are lying open on the table (summed up from different cards, or depicted on one single card), children must react as fast as possible and tap onto the card with the glass of marmalade on. The fastest child wins all the open cards on the table. Whenever a child taps onto the glass of marmalade and there are not five insects of the same kind, the child must distribute her/his cards amongst the other children. Whoever has all cards at the end wins the round. Level of difficulty was increased by adding inhibition (e.g., add cards where it is forbidden to react) or switching rules (e.g., it is only allowed to react to specific insects).

The game "*cheeky monkey*" taps selective attention, updating, visual search and speed of reaction; it was derived from a visual search task (Manly et al., 2001; Steele, Karmiloff-Smith, Cornish, & Scerif, 2012). A stack of cards depicting 6 items is placed in the middle of the table (face down), and a target card depicting 7 items is handed to each child. The target card depicts a tree house, with two children and their belongings on it. Each round, one card from the middle of the table (6 items) is turned around, and

children are instructed to find out which item is missing on this card compared to their target card (7 items). This item "had been stolen by the cheeky monkey". The player who identifies the missing item the fastest, gets the card. Whoever has the most cards in the end wins the round. Level of difficulty was increased by adding distractor items to the cards (more than six items are depicted on each card and they are not target items) and/or by taking the target card away, so that children had to remember each item from the target card (increasing interference control and working memory load).

The game "*medicine woman*" taps selective attention, cognitive flexibility, inhibition and updating; it is adapted from the "*keep track*" task (Van der Ven, Kroesbergen, Boom, & Leseman, 2012). In this game, one child is the game master and one is the player. After each round, they switch roles. The game master receives 12 different written instructions, and the player receives seven cards with depicted animals on them, along with a target card. The game master reads the instructions, and the player has to memorize names of rainforest animals that were either with the medical woman for treatment, or that could be released from her care. Subsequently, the child had to place the animals which can be released on the target card, and the ones which had to stay for treatment besides the game master. Level of difficulty was increased by a) the number of animals which had to be recalled, b) children had to remember the exact sequence in which the animals were released, and c) a rule switch was introduced for different kinds of animals, (e.g., that birds can fly home and therefore have not to be placed on the target card).

The game "*fruit salad*", taps inhibition, updating, switching and verbal fluency; it was derived from an updating and verbal fluency task (Archibald & Kerns, 1999; Munakata, Snyder, & Chatham, 2012). This game consists of 80 cards, with every card depicting one of four fruits (pineapple, banana, melon, orange). The cards were equally

distributed to the players, with each having a stack of cards in front of them. Players had to put one card each – one after the other – in the middle of the table. As fast as possible they must say out loud the name of the fruit depicted on their card. However, they also had to remember the card which was turned over before, because it was not allowed to name the same fruit twice in a row. Thus, when their depicted fruit was already turned over and named by the player beforehand (e.g. melon), players quickly had to come up with a different fruit name (e.g. apple). If then the next player turns over the same fruit again, he had to switch back again and name it by its real name (e.g. melon). Level of difficulty was increased by adding several inhibition cards (depicting a rotten version of each fruit). Whenever a player turns over one of these card, it was forbidden to name the depicted rotten fruit by its real name. Instead, every time the player turned over this fruit (e.g. rotten banana), players had to come up with a different fruit. The rotten fruit cards were either lying face down or face up on the table (increasing the working memory demands).

Outcome measures

EF were assessed by means of computer tasks using E-Prime Software (Psychology Software Tools, Pittsburgh, PA). In several studies with children and adolescents, these tasks have been proven to be reliable and valid measures of EF (Jäger, Schmidt, Conzelmann, & Roebers, 2014, 2015; Schmidt, Jäger, Egger, Roebers, & Conzelmann, 2015).

Updating was assessed by means of a non-spatial visual 2-back task, which was adapted from a spatial n-back task developed by Chen, Mitra, and Schlaghecken (2008). In this task, pictures of fruits are displayed on the screen one after another. In cases where the fruit on the screen corresponded to the preceding fruit presented, children had to press the right button. In all other cases, the left button had to be pressed. The

task consisted of two test blocks containing 24 trials each, with one third of trials being targets. The total number of correct answers was used as the dependent measure.

Inhibition was measured by means of a child-adapted Flanker task (Roebers & Kauer, 2009). In this task, five red fish were presented on a screen. The children were instructed to feed the fish by indicating in which direction the fish in the middle had his mouth, by pressing an external response button. The fish in the middle could either swim in the same direction as the four-flanking fish (congruent trials), or in the opposite direction (incongruent trials). It consisted of a block with 20 congruent trials ("pure" block), and a block with a total of 40 trials; that is, 20 congruent and 20 incongruent trials in a randomized order ("standard" block). Inter-stimuli- intervals varied randomly from 800 to 1400 ms. Mean reaction times of the conflict scores (Rueda, Posner, & Rothbart, 2005) between the incongruent trials of the standard block (trials with the highest rate of distraction) and the congruent trials of the pure block (trials with the lowest rate of distraction) were calculated as the dependent measure of inhibition.

Shifting was assessed by an additional block ("mixed" block), presented in the same graphic design as the Flanker task (Röthlisberger et al., 2012). In this block, again, 20 congruent and 20 incongruent trials were shown with an additional rule being introduced, cued by different trial-colours. In the red fish trials, the fish in the middle was the target fish; for the yellow fish, it was the four flanking fish. Children were instructed to feed the fish in the middle for the red fish, and the four-flanking fish for the yellow fish. Again, the fish in the middle and the four-flanking fish could either swim in the same (congruent trials) or in the opposite direction (incongruent trials). Whenever children had to adapt their response depending on the colour of the stimuli, a switch between both rules was required (switching trials). Inter-stimuli- intervals varied randomly from 800 to 1400 ms. Mean reaction times of global switch costs

between the shifting and inhibition block were calculated as the dependent variable (Chevalier & Blaye, 2009). Since trials in the mixed block not only required the child to shift between different tasks but also contained inhibitory demands, the difference between mixed and standard block was calculated to control for the inhibition component.

Statistical analyses

Statistical tests were performed using SPSS 23.0 (SPSS Inc., Chicago, IL, USA). As *outlier analysis,* according to previous studies using the same measures trials with a reaction time under 150 ms were excluded as anticipatory responses (interindividual outliers < 0.5%). In a next step, trials with reaction times deviating by more than 3 SD from the child's mean (intraindividual outliers < 1%) were excluded as well, because they might be influenced by other factors such as difficulty staying on task. Only correct trials were included in the calculation of reaction times. Subsequently, blocks with an accuracy of less than or equal to 50% were deleted, assuming that participants did not understand the task correctly in this case (less than 2%). Pattern of results did not change when outliers were excluded. As *multivariate outlier analyses*, the Mahalanobis distance values (Fidell & Tabachnick, 2003) were calculated as χ^2 at $p < .001$ with 6 degrees of freedom (equal to the number of variables updating, inhibition and shifting at pre- and post-test). One case was identified as a probable outlier, and excluded having a Mahalanobis distance greater than 22.46. The exclusion did not affect pattern of results.

Preliminary analyses were performed using *t*-tests for between group comparisons of background variables (see Table 1) and dependent variables at pre-test. No significant differences could be observed in both (ps > .05).

For *main analyses,* one-tailed ANCOVAs, using pre-test performance as

covariates and post-test performance as dependent variable, were conducted separately for each computerized test. In addition, results of two-tailed tests will be reported. Partial eta square (η^2_p) was reported as an estimation of effect size. The level of significance was set at $p < .05$ for all analyses. To examine potential differential effects for *exploratory analyses,* the change in the three core EF (updating, inhibition, shifting) between pre- and post-test were correlated with background variables.

Results

In terms of updating ($F(1,115) = 3.42$, $p = .034$, $\eta^2_p = .029$) and shifting ($F(1,115) = 7.54$, $p = .004$, $\eta^2_p = .062$), the performance significantly improved in the cognitive games condition compared to control (see Table 2, Fig. 1). In contrast, no significant differences were found for inhibition ($F(1,115) = 0.45$, $p = .252$, $\eta^2_p = .004$). When considering two-tailed analyses, there was a trend for updating ($p = .068$) and a significant effect for shifting ($p = .008$). In summary, these results indicate that the cognitive games condition improved specific EF (updating and shifting), showing a stronger effect on shifting than updating. Regarding potential differential effects, gain scores were compared against background variables. No significant correlations (ps $> .05$) were found, indicating that participants could benefit irrespective of age, gender and pubertal status at the intervention.

Discussion

In the current study, cognitive training in small groups carried out by class teachers improved performance of untrained shifting and updating tasks in older primary school children (ages 10-12). The customized, adaptive, school-based training was designed to directly and indirectly target EF, resulting in enhancement after only six weeks of

training.

The main finding of this study is that cognitive training promotes EF (updating and shifting) in older primary school children; revealing a stronger effect for the shifting component than for updating. It has been argued that improving EF early in life is critical, and that training these higher-cognitive processes has greater effects in younger children (Diamond, 2016). Previously, most school-based interventions have been developed to suit the age range of 4 to 10 years, however, results of this study revealed that school-based interventions can also contribute to positive EF development in older primary school children (ages 10-12). This is important, not only when considering poor performance in EF among all age-populations in school children (Riccio & Gomes, 2013; Zelazo & Müller, 2010), but also for the crucial role EF seems to play in the transition to middle school (Jacobson et al., 2011). The findings of the current study are in line with several studies investigating cognitive training, indicating that EF can be improved at any age (Diamond & Ling, in press).

In this study, significant effects on updating and shifting were observed. These results are partly in accordance with a previous study, which found positive effects on all three core EF after six weeks of cognitive training in pre-school children (Röthlisberger et al., 2012). The intervention adapted for older ages was found to be feasible and effective in older primary school children, with regard to updating and shifting. Interestingly, the positive trend effect observed on updating was not significant after two-tailed testing, nevertheless, a positive effect on core EF is highly relevant, since they exhibit strong associations with academic achievement. More precisely, shifting performance predicts school performance in mathematics and reading (Yeniad, Malda, Mesman, van Ijzendoorn, & Pieper, 2013); and updating is consistently associated with mathematics, literacy, and second-language learning (Latzman et al.,

2010; Lee et al., 2012; St Clair-Thompson & Gathercole, 2006). To examine the interaction between each core EF and their associated academic areas, future studies may want to focus on interventions that specifically train only one EF, but examine its effects on all three core EF and subsequently on academic achievement.

For inhibition, no effect was found in the current study. In general, the empirical evidence regarding positive effects on inhibition is contradictory (Diamond & Ling, in press). While computerized approaches targeting inhibition failed to find positive effects (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; Thorell et al., 2009), more general interventions such as specific school curricula and programs on inhibition showed beneficial effects (Blair & Raver, 2014; Diamond, Barnett, Thomas, & Munro, 2007). Hence, this indicates that improving inhibition in later school years is possible, but might be less easily achieved. This interpretation is supported by the developmental trajectories of EF, showing that inhibition is the first to fully mature, and therefore might be more stable and less easily affected in older compared to younger children (Diamond, 2013; Jäger et al., 2014). Therefore, an important issue for further research is to tailor an appropriate intervention that improves EF by considering the child's developmental EF timetable.

Differential effects and existing heterogeneity in certain demographic variables such as age and pubertal status raise the question whether all children benefit equally from an applied intervention. However, since no moderating variables were found, we are tempted to believe that children profited equally, irrespective of their age, gender and pubertal status. Considering that there are always children who apply more effort to an intervention, we explored gain score distributions on a descriptive level, and thus observed hints for a more pronounced improvement in children with lower EF values at baseline. This is in line with several reviews indicating that children with poor

baseline performance benefit most, which may be because there is more room for improvement (Diamond & Ling, 2016; Otero et al., 2014). In general, little is known about variables that distinguish children who benefit from those who do not. This is an issue that should be examined by future studies.

In this study, a gamified cognitive group training was carried out, targeting multiple EF directly and indirectly. Although it has been recently claimed that multimodal training approaches might be more effective than one-dimensional approaches, such as computerized training (Cardoso et al., 2016), most studies using school-based interventions targeted only one core EF domain (Cardoso et al., 2016; Diamond & Lee, 2011; Otero et al., 2014). While direct trainings might promote larger training gains and smaller transfer effects, indirect trainings seem to have opposite effects (Diamond & Ling, in press). Given this trade-off, the current study combined direct and indirect trainings targeting multiple EF. The administered cognitive games specifically targeted core EF components, and successful game play, and thus efficient training, was dependent on social interaction including emotion regulation and social competencies. Although the combination of direct and indirect means constitutes a strength of the present study, the extent to which their combination affected transfer effects remains unclear.

For shifting and updating, positive effects to untrained computerized EF task performance were observed. These results seem to be in line with a study by Mackey et al. (2011), in which the authors predicted larger gains in their interventional approach (compared to previous studies), because they included direct and indirect factors in their training. In fact, they observed positive effects in untrained (intervention related) tests on fluid reasoning and processing speed, by letting children (aged 7 to 9) play commercially available games in an after-school program. Thus, one might assume that

direct and indirect factors might have beneficially influenced transfer effects in both studies. However, the experimental designs applied could not separate the impacts of direct and indirect factors that led to improved EF performance. Future studies should compare integrated approaches with direct and indirect trainings to disentangle the single and combined effects of cognitive training on EF performance.

This present study does not come without limitations. First, the experimental group was compared to a control group which had regular school lessons instead of the gamified cognitive training. Although both groups were directed by their regular class teachers, a potential improvement due to the additional attention required (the so-called 'Hawthorne effect') is a major limitation of the current study. Thus, alternative explanations to an improvement in EF performance, such as an increase in motivation, or familiarization with strategies regarding these types of tasks, might also apply. This is of particular importance with regard to recent reviews demonstrating that effect sizes were significantly reduced when active control groups were considered as a comparison (Melby-Lervag & Hulme, 2013; Shipstead et al., 2012). Second, based on the literature concerning the relation between EF and academic achievement, we assumed a positive relationship between the two. However, neither a far transfer testing (such as measuring academic achievement) nor a follow-up assessment was conducted in this study. This is an important issue considering the current controversial debate of whether training effects also transfer to other domains (Melby-Lervag & Hulme, 2013). Future research is needed to investigate this. Third, cluster randomization instead of individual randomization was used. For ecological validity and ethical reasons, it was not possible to split up the different classes and randomize on an individual level. Consequently, because of a lack of independence within one cluster, complexity was possibly introduced into the design and analysis. To account for this, however, we included a

larger sample size (Donner, Birkett, & Buck, 1981; Rutterford, Copas, & Eldridge, 2015). Fourth, another limitation is that each of the three core EF was assessed by only one task. Considering the task-impurity problem and that any core EF is measured by a task including not only variance of EF processes but also non-EF processes, using multiple tasks per EF component would have been advantageous (Miyake & Friedman, 2012). Consequently, in the current study also other task specific non-EF processes could have had an impact on the obtained effects. Unfortunately, it was not feasible for this study, as the assessment would have been too long for children, and would have increased the risk of washing out potential intervention effects resulting from multiple tests and fatigue.

Conclusion

In summary, this study establishes an important finding regarding school-based cognitive training, demonstrating that a small-group intervention targeting core EF is feasible and effective in older primary schoolchildren. In contrast to individualized time- and cost-consuming interventions, this small-group intervention is highly scalable and can easily be carried out in school lessons. Our findings are therefore of practical relevance in the field of education. Further and regular implementation of such EF-fostering activities into children`s school routines may constitute an aim for educational practice. Exploring the long-term impact on academic achievement and mental health, and to investigate the sustainability of effects, should be worthwhile issues for future research.

References

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*(2), 115–129. http://doi.org/10.1076/chin.5.2.115.3167

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. http://doi.org/10.1016/j.lindif.2011.01.007

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology*, *20*(3), 899–911. http://doi.org/10.1017/S0954579408000436

Blair, C., & Raver, C. C. (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. *PLoS ONE*, *9*(11). http://doi.org/10.1371/journal.pone.0112393

Cardoso, C. O., Dias, N., Senger, J., Colling, A. P. C., Seabra, A. G., & Fonseca, R. P. (2016). Neuropsychological stimulation of executive functions in children with typical development: A systematic review. *Applied Neuropsychology: Child*, *0*(0), 1–21. http://doi.org/10.1080/21622965.2016.1241950

Chen, Y.-N., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the N-back task: An investigation using ERPs. *Clinical Neurophysiology*, *119*(7), 1546–1559. http://doi.org/10.1016/j.clinph.2008.03.003

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. http://doi.org/10.1016/j.neuropsychologia.2006.02.006

Diamond, A. (2010). The evidence base for improving school outcomes by addressing the whole child and by addressing skills and attitudes, not just content. *Early Education and Development*, *21*(5), 780–793. http://doi.org/10.1080/10409289.2010.514522

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

Diamond, A. (2016). Why improving and assessing executive functions early in life is critical. In J. S. Griffin, P. McCardle & L. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 11–43). Washington: American Psychological Association. http://doi.org/10.1037/14797-002

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*(5855), 1387–1388. http://doi.org/10.1126/science.1151148

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

Diamond, A., & Ling, D. S. (in press). Review of the evidence on, and fundamental questions surrounding, efforts to improve executive functions. In M. Bunting, J. Novick, M. Dougherty, & R. W. Engle (Eds.), *An integrative approach to cognitive and working memory training: Perspectives from psychology,*

neuroscience, and human development. Oxford, UK: Oxford University Press.

Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, *18*, 34–48. http://doi.org/10.1016/j.dcn.2015.11.005

Donner, A., Birkett, N., & Buck, C. (1981). Randomization by cluster. Sample size requirements and analysis. *American Journal of Epidemiology*, *114(6)*(6), 906–914.

Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A. C., & Moehr, M. J. (2011). Executive control and dimensions of problem behaviors in preschool children. *Journal of Child Psychology and Psychiatry*, 52(1), 33–46.

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., Corley, R. P., & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science*, *18*(10), 893–900. http://doi.org/10.1111/j.1467-9280.2007.01997.x

Gooch, D., Thompson, P., Nash, H. M., Snowling, M. J., & Hulme, C. (2016). The development of executive function and language skills in the early school years. Journal of Child Psychology and Psychiatry, 57(2), 180–187. http://doi.org/10.1111/jcpp.12458

Jacobson, L. A., Williford, A. P., & Pianta, R. C. (2011). The role of executive function in children's competent adjustment to middle school. *Child Neuropsychology*, *17*(3), 255–280. http://doi.org/10.1080/09297049.2010.535654

Jäger, K., Schmidt, M., Conzelmann, A., & Roebers, C. M. (2014). Cognitive and physiological effects of an acute physical activity intervention in elementary

school children. *Frontiers in Psychology*, *5*, 1473.
http://doi.org/10.3389/fpsyg.2014.01473

Jäger, K., Schmidt, M., Conzelmann, A., & Roebers, C. M. (2015). The effects of qualitatively different acute physical activity interventions in real-world settings on executive functions in preadolescent children. *Mental Health and Physical Activity*, *9*, 1–9. http://doi.org/10.1016/j.mhpa.2015.05.002

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology*, *32*(5), 455–462. http://doi.org/10.1080/13803390903164363

Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., & Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: Executive functioning, patterns, numeracy, and arithmetic. *British Journal of Educational Psychology*, *82*(1), 82–99. http://doi.org/10.1111/j.2044-8279.2010.02016.x

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science*, *14*(3), 582–590. http://doi.org/10.1111/j.1467-7687.2010.01005.x

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A, Watson, P., & Robertson, I. H. (2001). The differential assessment of children's attention: the Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(8), 1065–1081. http://doi.org/10.1111/1469-7610.00806

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291. http://doi.org/10.1037/a0028228

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual

differences in executive functions. *Current Directions in Psychological Science*,

21(1), 8–14. http://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager,

T. D. (2000). The unity and diversity of executive functions and their

contributions to complex "frontal lobe" tasks: a latent variable analysis.

Cognitive Psychology, *41*(1), 49–100. http://doi.org/10.1006/cogp.1999.0734

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive

control: three key transitions. *Current Directions in Psychological Science*,

21(2), 71–77. http://doi.org/10.1177/0963721412436807

Neuenschwander, R., Röthlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How

do different aspects of self-regulation predict successful adaptation to school?

Journal of Experimental Child Psychology, *113*(3), 353–371.

http://doi.org/10.1016/j.jecp.2012.07.004

Otero, T. M., Barker, L. a, & Naglieri, J. a. (2014). Executive function treatment and

intervention in schools. *Applied Neuropsychology: Child*, *3*(3), 205–214.

http://doi.org/10.1080/21622965.2014.897903

Pellegrini, A. D., Blatchford, P., Kato, K., & Baines, E. (2004). A Short-term

Longitudinal Study of Children's Playground Games in Primary School:

Implications for Adjustment to School and Social Adjustment in the USA and

the UK. *Social Development*, *13*(1), 107–123. http://doi.org/10.1111/j.1467-

9507.2004.00259.x

Riccio, C. A., & Gomes, H. (2013). Interventions for executive function deficits in

children and adolescents. *Applied Neuropsychology: Child*, *2*(2), 133–140.

http://doi.org/10.1080/21622965.2013.748383

Roebers, C. M. (2017). Executive functions and metacognition: towards a unifying framework of cognitive self-regulation. *Developmental Review,* 45, 31-51. http://dx.doi.org/10.1016/j.dr.2017.04.001

Roebers, C. M., & Kauer, M. (2009). Motor and cognitive control in a normative sample of 7-year-olds. *Developmental Science*, *12*(1), 175–181. http://doi.org/10.1111/j.1467-7687.2008.00755.x

Roebers, C. M., Röthlisberger, M., Neuenschwander, R. & Cimeli, P. (2014). Nele und Noa im Regenwald: Berner Material zur Förderung exekutiver Funktionen – Spielebox [Nele and Noa in the rainforest: Bernese material to promote exekutive functions – gamebox]. Reinhardt Verlag München.

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. http://doi.org/10.1002/icd.752

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention M. *Proceedings of the National Academy of Sciences*, *102*(41), 14931–14936. http://doi.org/10.1073/pnas.0506897102

Rutterford, C., Copas, A., & Eldridge, S. (2015). Methods for sample size determination in cluster randomized trials. *International Journal of Epidemiology*, *44*(3), 1051–1067. http://doi.org/10.1093/ije/dyv113

Schmidt, M., Jäger, K., Egger, F., Roebers, C. M., & Conzelmann, A. (2015). Cognitively engaging chronic physical activity, but not aerobic exercise, affects executive functions in primary school children: A group-randomized controlled trial. *Journal of Sport and Exercise Psychology*, *37*(6), 575–591.

http://doi.org/10.1123/jsep.2015-0069

Shipstead, Z., Hicks, K. L., & Engle, R. W. (2012). Cogmed working memory training: Does the evidence support the claims? *Journal of Applied Research in Memory and Cognition*, *1*(3), 185–193. http://doi.org/10.1016/j.jarmac.2012.06.003

Sjöwall, D., Bohlin, G., Rydell, A.-M., & Thorell, L. B. (2017). Neuropsychological deficits in preschool as predictors of ADHD symptoms and academic achievement in late adolescence. *Child Neuropsychology*, *23*(1), 111–128. http://doi.org/10.1080/09297049.2015.1063595

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. http://doi.org/10.1016/j.cogdev.2016.02.003

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*(4), 745–759. http://doi.org/10.1080/17470210500162854

Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The Multiple Subfunctions of Attention: Differential Developmental Gateways to Literacy and Numeracy. *Child Development*, *83*(6), 2028–2041. http://doi.org/10.1111/j.1467-8624.2012.01809.x

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

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*(1), 100–119. http://doi.org/10.1111/j.2044-8279.2011.02035.x

Vitaro, F., Brendgen, M., Larose, S., & Trembaly, R. E. (2005). Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of Educational Psychology*, *97*(4), 617–629. http://doi.org/10.1037/0022-0663.97.4.617

Viterbori, P., Usai, M. C., Traverso, L., & De Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in Grades 1 and 3: A longitudinal study. *Journal of Experimental Child Psychology*, *140*, 38–55. http://doi.org/10.1016/j.jecp.2015.06.014

Watzlawik, M. (2009). Die Erfassung des Pubertätsstatus anhand der Pubertal Development Scale: Erste Schritte zur Evaluation einer deutschen Übersetzung [Assessing pubertal status with the Pubertal Development Scale: First steps towards an evaluation of a German translation]. *Diagnostica*, *55*(1), 55–65. http://doi.org/10.1026/0012-1924.55.1.55

Yeniad, N., Malda, M., Mesman, J., van Ijzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1–9. http://doi.org/10.1016/j.lindif.2012.10.004

Zelazo, P. D., & Müller, U. (2010). Executive function in typical and atypical development. In U. Goswami (Eds.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development* (pp. 574–603). Oxford, UK: Wiley-Blackwell. http://doi.org/10.1002/9781444325485.ch22

Tables

Table 1. Means, standard deviations and test statistics comparing background variables between groups.

	Cognitive games ($n = 61$)	Control ($n = 57$)		
			t	p
	M (SD)	M (SD)	(116)	
Age (years)	11.31 (0.62)	11.48 (0.73)	- 1.33	.188
Sex (% female)	48.3	51.7	0.01	.995
Body mass index (kg/m^2)	18.32 (2.97)	17.74 (2.73)	1.10	.275
Pubertal status	4.84 (1.45)	5.04 (1.52)	-0.72	.472

**Table 2.** Executive function performance before and after intervention.

	Cognitive games ($n = 61$)	Control ($n = 57$)
	M (SD)	M (SD)
Pre-test		
Updating (accuracy)[a]	10.82 (3.09)	11.18 (2.56)
Inhibition (RT)[b]	100.06 (83.92)	81.23 (67.75)
Shifting (RT)[b]	597.90 (216.24)	565.49 (203.99)
Post-test		
Updating (accuracy)[a]	12.27 (2.72)	11.46 (3.53)
Inhibition (RT)[b]	69.88 (64.19)	68.39 (68.05)
Shifting (RT)[b]	401.44 (160.47)	457.00 (157.16)
Pre-to-posttest Data		
Δ Updating (accuracy)[a]*	1.46 (3.23)	0.28 (3.20)
Δ Inhibition (RT)[b]	-30.17 (75.12)	-12.84 (64.31)
Δ Shifting (RT)[b]*	-196.46 (158.80)	-108.49 (208.17)

Note. [a]*Accuracy corresponds to the number of correct responses.*

[b]*Reaction times are presented in milliseconds.*

* p < .05.

1   Figures
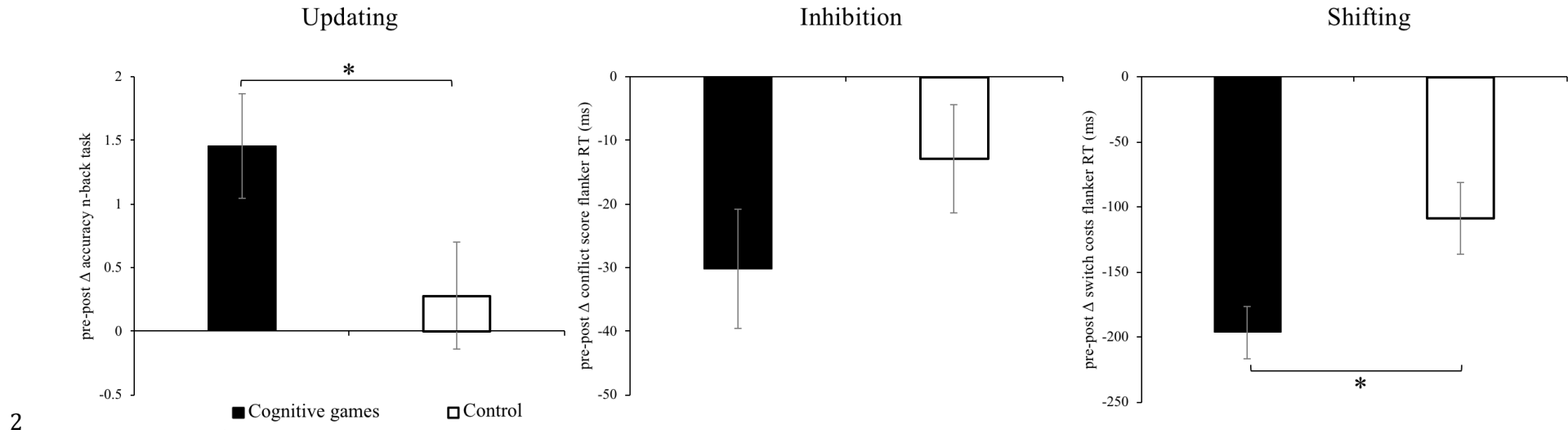
2

3   **Figure 1.** Means and error bars (representing the standard error of the mean) for the change (Δ) in the three core EF (updating, inhibition and

4   shifting) in the two experimental conditions between pre- and post-test. RT = reaction time. * $p < .05$.