Journal of Experimental Psychology:
Learning, Memory, and Cognition
2014, Vol. 40, No. 4, 987–1001

© 2014 American Psychological Association
0278-7393/14/$12.00 DOI: 10.1037/a0036165

No Evidence for True Training and Transfer Effects After Inhibitory Control Training in Young Healthy Adults

Sören Enge, Alexander Behnke,
Monika Fleischhauer, and Lena Küttler
Technische Universitaet Dresden

Matthias Kliegel
University of Geneva

Alexander Strobel
Technische Universitaet Dresden

Recent studies reported that training of working memory may improve performance in the trained function and beyond. Other executive functions, however, have been rarely or not yet systematically examined. The aim of this study was to test the effectiveness of inhibitory control (IC) training to produce true training-related function improvements in a sample of 122 healthy adults using a randomized, double-blind pretest/posttest/follow-up design. Two groups performed either adaptive (training group) or nonadaptive (active control) versions of go/no-go and stop-signal tasks for 3 weeks. Training gains as well as near-transfer to an untrained Stroop task and far-transfer to psychometric fluid intelligence were explored. Although the adaptive group could substantially improve overall IC task performance after training, no differences to the active control group occurred, neither at posttest nor at follow-up testing. A large decrease in response latency from pre- to posttest (and from pretest to 4 months' follow-up testing) was found when the training group was compared to the passive control group, which, however, does not sufficiently control for possible confounds. Thus, no conclusive evidence was found that this performance increase mirrors a true increase in IC function. The fact that training improvement was mainly related to response latency may indicate that individuals were more focused on performance gains in the prepotent go trials but less on the stop trials to meet the requirements of the tasks as well as possible. The challenges for response inhibition training studies are extensively discussed.

Keywords: inhibitory control, response inhibition, training, transfer

A number of recent studies reported training and transfer effects after working memory (WM) training in healthy individuals and in samples with executive function impairments (see, e.g., Klingberg, 2010; Morrison & Chein, 2011). For instance, Klingberg, Forssberg, and Westerberg (2002) reported that children with attention-deficit/hyperactivity disorder (ADHD) who practiced on a visuospatial and verbal WM training not only improved on the trained tasks, but also showed transfer on Stroop performance and on psychometric fluid intelligence (Gf). Further data suggest that untrained executive measures and psychometric Gf may be sensitive to WM training: Jaeggi, Buschkuehl, Jonides, and Perrig (2008), for example, reported far-transfer on Gf in a healthy sample attending to an adaptive WM-updating training using a dual *n*-back task (see also Jaeggi et al., 2010).

However, given the considerable heritability estimates for Gf (40%–80%) emphasizing the importance of genetic variance in establishing individual differences in intelligence (e.g., Plomin & Spinath, 2002), a real improvement of Gf by a simple training of WM functions would be astonishing (Moody, 2009; Shipstead, Redick, & Engle, 2010) but may support the notion that Gf is not invariant to changes (Sternberg, 2008). Moreover, converging evidence suggests a strong relationship between WM measures and inductive reasoning, with correlations of up to .90 (Buehner, Krumm, & Pick, 2005; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Kyllonen & Christal, 1990). Thus, WM appears to be (partly) a possible basis for Gf (Sternberg, 2008), rendering it reasonable that training of the former may impact on the latter. However, other studies using the cognitive training program (Cogmed Systems AB) developed by Klingberg and colleagues (see Klingberg et al., 2005) provided inconsistent results about the effectiveness of cognitive training: For preschool children with ADHD, Klingberg et al. (2005) could replicate the WM training-related improvements in measures of inhibitory control (IC; Stroop task) and reasoning (Raven's Colored Progressive Matrices), as previously reported (Klingberg et al., 2002), and albeit in a sample of only three participants, found increased functional magnetic

---

Sören Enge, Alexander Behnke, Monika Fleischhauer, and Lena Küttler, Department of Psychology, Technische Universitaet Dresden; Matthias Kliegel, Department of Psychology, University of Geneva; Alexander Strobel, Department of Psychology, Technische Universitaet Dresden.

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 940/1 2012). We are indebted to Anne Gärtner for her valuable help in data acquisition and Katharina Zinke for advice during the planning phase of this study. We are also indebted to all of the participants for their time and effort.

Correspondence concerning this article should be addressed to Sören Enge, Department of Psychology, Technische Universitaet Dresden, TU Dresden, 01062, Germany. E-mail: enge@psychologie.tu-dresden.de

resonance imaging (MRI) BOLD activation in frontal and parietal regions as a concomitant of training (Olesen, Westerberg, & Klingberg, 2004; see also Westerberg & Klingberg, 2007; Takeuchi et al., 2010). Further attempts using the Cogmed program were unsuccessful in replicating transfer on untrained IC or Gf tasks, including samples of healthy preschool children (Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009), children with low WM (Holmes, Gathercole, & Dunning, 2009; van der Molen, 2010), children with ADHD (Holmes et al., 2010), and adult patients suffering from stroke-related impairments (Westerberg et al., 2007). Consistently, using a spatial and verbal WM training, Chein and Morrison (2010) reported an increase for Stroop performance after training, but not for reasoning (Advanced Progressive Matrices [APM]). Nonetheless, training gains that have been observed in these studies suggests that improvements in the trained task may be preserved over time and that near-transfer may also be possible, whereas no conclusive evidence has been provided for far-transfer (Enriquez-Geppert, Huster, & Herrmann, 2013; Klingberg, 2010; Morrison & Chein, 2011; Redick et al., 2013; Shipstead et al., 2010; Shipstead, Redick, & Engle, 2012).

Given that the improvement of executive functions is of high individual and social significance and would provide therapeutic potential for cognitive deficits, training gains that can be attributed to real changes in the trained function would be of particular value (e.g., Klingberg, 2010). Thus, more data about the temporal stability and the size of putative training effects would be desirable. This holds particularly true for the effectiveness of training of other important functions beside WM, for which no or only insufficient data are currently available (Klingberg, 2010; Shipstead et al., 2010).

In this regard, although there is no consensus about the exact number, classification, and relationship among the postulated executive functions so far, inhibition or IC has been widely considered to be one of the important factors in cognitive and behavioral regulation (Aron, Robbins, & Poldrack, 2004; Diamond, 2013; Friedman & Miyake, 2004; Miyake et al., 2000; Salthouse, Atkinson, & Berish, 2003). Further, systematic examinations of explained variances in cognitive tasks have demonstrated that IC can affect task performance incrementally over WM and task shifting (Engle, Tuholski, Laughlin, & Conway, 1999; Fisk & Sharp, 2004; Miyake et al., 2000; but see Friedman et al., 2006). Basically, IC can be defined as the ability to suppress stimuli, prepotent responses or impulses, behavioral alternatives, overlearned habits, interpretations, and memories that are currently irrelevant, interfering, incorrect, or inappropriate to perform goal-directed behavior (Aron et al., 2004; Barkley, 1997; Ridderinkhof, van den Wildenberg, Segalowitz, & Carter, 2004; Salthouse et al., 2003).

Consequently, IC is crucial to focus on relevant information in complex environments (Barkley, 1997), playing a role in verbal communication, reading comprehension (De Beni, Palladino, Pazzaglia, & Cornoldi, 1998), and memory retrieval, mediating the suppression of inappropriate thoughts (Anderson & Weaver, 2009; Levy & Anderson, 2002). Moreover, IC affects planning (Barkley, 1997) and prospective memory (Kliegel, Mackinlay, & Jager, 2008), promotes decision making and cognitive flexibility (Bechara, 2005; Clark, Cools, & Robbins, 2004), and supports learning and problem solving (Dempster & Corkill, 1999b; Passolunghi, Cornoldi, & De Liberto, 1999), as demonstrated for mathematical aptitude (Dempster & Corkill, 1999a), and by associations with Gf tasks up to $r = .73$ (Salthouse et al., 2003). Thus, IC can be considered as a fundamental function for "successful living" (Garavan, Ross, & Stein, 1999, p. 8301).

On the contrary, IC deficits are commonly associated with several clinical syndromes, as primarily with ADHD (Barkley, 1997), but also with schizophrenia (Nestor & O'Donnell, 1998), autism (Ciesielski & Harris, 1997), obsessive-compulsive disorder (Bannon, Gonsalvez, Croft, & Boyce, 2002; Enright & Beech, 1993), and drug addiction (Bechara, 2005; Giancola & Tarter, 1999; Perry & Carroll, 2008), and have been suggested to be one of the key developmental mechanisms for cognitive aging (Diamond, 2013; Hasher & Zacks, 1988).

With regard to training studies, near-transfer of WM training on IC-related paradigms like the Stroop were reported for individuals with executive function deficits such as children with ADHD and older adults with age-related impairments in executive control functions, as well as for healthy populations (for review, see Enriquez-Geppert et al., 2013; Klingberg, 2010; Morrison & Chein, 2011; Shipstead et al., 2012): For example, near-transfer on Stroop performance was reported for children with ADHD using the Cogmed program (Klingberg et al., 2002), which remained stable even 3 months after the training (Klingberg et al., 2005). Further, transfer on Stroop performance was reported for a spatial and verbal WM training in young adults (Chein & Morrison, 2010), and for a task-shifting training in different age groups (Karbach & Kray, 2009). In contrast, Thorell et al. (2009) found no near-transfer of WM training on Stroop performance in healthy preschool children, but limited transfer on performance in the go/no-go task (omission errors). No transfer effects occurred when IC was trained directly. Owens, Koster, and Derakshan (2013) showed that dual *n*-back training resulted in improved visual IC in dysphoric participants. In old-old individuals, Borella, Carretti, Zanoni, Zavagnin, and De Beni (2013) observed successful transfer of WM training on Stroop performance.

All in all, it is notable that most of the available studies using IC tasks have been based on populations with executive deficits, children or older adults; used relatively small samples; yielded inconsistent results; and in most cases served to test potential transfer effects of WM training. Moreover, conclusions on the effectiveness of cognitive training were frequently drawn by comparing training with passive control groups, which, however, may not sufficiently control for potential confounds, as the groups differ in exposure to tasks and laboratory conditions (Redick et al., 2013; Shipstead et al., 2012).

Taken together, a systematic evaluation of the temporal stability and the effect size of IC training (and possible transfer) is lacking so far. Thus, the main aim of this study was to provide an estimation of the effectiveness of IC training to produce true training effects in a young and healthy adult population. For this purpose, the go/no-go and stop-signal tasks were used as training tasks. These well-established and among the most widely used IC tasks measure (motor) response inhibition as a major form of IC in a relatively pure manner (see, e.g., Aron et al., 2004; Logan, 1994; Miyake et al., 2000; Ridderinkhof et al., 2004; Simmonds, Pekar, & Mostofsky, 2008). Further, it was of interest whether IC training caused near-transfer on the performance of an untrained Stroop task that is considered a valid and more complex marker of IC (Friedman & Miyake, 2004; MacLeod, 1991; Ridderinkhof et al., 2004; Salthouse et al., 2003; Simmonds et al., 2008). For reasons of comparability with other studies and because IC has been related to Gf measures (Dempster & Corkill, 1999b; Engle et al., 1999; Fisk & Sharp, 2004; Kane & Engle, 2002), far-transfer of IC training on psychometric Gf was also explored.

Considering methodological advice (Klingberg, 2010; Moody, 2009; Shipstead et al., 2010, 2012; Sternberg, 2008), the study was designed as a pretest, posttest, and follow-up trial in a sample of 122 participants and comprised two groups that performed either adaptive (training group) or nonadaptive (active controls) versions of the two training tasks during 3 weeks. The group assignment was randomized and double-blind. Both groups were compared to test the effectiveness of IC training. In addition, a passive control group was implemented to provide comparison with previous studies relying on no-contact controls.

Method

Participants

The present study was conducted in accordance with the Declaration of Helsinki and followed the ethical standards of the German Psychological Association. The total sample comprised 122 student volunteers (105 women, 17 men; mean age = 21.3 years, *SD* = 4.16; range: 18–38) who gave written informed consent prior to the beginning of the study. Individuals received course credit for their participation and were fully debriefed upon completion of the study. All participants were randomly assigned to either one of three groups: (a) the adaptive training group (*n* = 43), where task difficulty was individually adjusted on a trial-by-trial basis; (b) the nonadaptive group (*n* = 39), which served as active control group, as task difficulty was fixed according to the individual performance level obtained at the pretest session (i.e., participants performed far below their capacity limit); or (c) the passive control group (*n* = 40), which did not receive training or any other task, but was implemented to provide comparison with other studies and to control for retest effects. There was a dropout of two individuals in the posttest session and of nine in the follow-up session 4 months after the training. All subjects had normal or corrected-to-normal vision, and self-report data revealed no history of relevant health problems.

Experimental Design

Each testing (pretest, posttest, follow-up) lasted about 2 hr and consisted of two blocks. First, Gf and questionnaires measuring mood and the possible confounding factors sex, age, malaise, sleep duration, smoking, and caffeine and alcohol consumption were assessed. Second, IC tasks were conducted: Participants performed a Stroop task (10 min), a stop-signal task (15 min), and a go/no-go task (15 min) with counterbalanced task order.

After pretesting, participants were randomly assigned to one of the three experimental conditions by an uninvolved person and a coding scheme, which assured that neither experimenters nor participants were informed of the individual's group assignment. During the 3 weeks between pre- and posttesting, participants of the adaptive training group and the active control group completed three training sessions per week. It was ensured that participants did not train on three consecutive days to prevent possible effects of massed learning, as spaced or distributed learning has been shown to be more effective (for review, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). At the beginning of each session, participants rated their current mood. Subsequently, two blocks of 15 min each followed in which the computerized go/no-go and stop-signal training tasks were performed (see below). The order of the training tasks was counterbalanced over the nine sessions. After the 3 weeks' training period, all participants returned for posttesting and about 4 months later for the follow-up assessment.

Measures

Inhibitory control tasks. Three IC tasks used either as training tasks (go/no-go and stop signal) or as near-transfer task (Stroop) were administered using the Presentation software (http://www.neurobs.com) running at LCD screens with a resolution of 1024 × 768. The participants read a written instruction and practiced each task for at least 20 trials.

Go/no-go task (Training Task 1). In the go/no-go task (see Figure 1A), a sequence of black capital letters (*A*–*Z*) in 72-point Arial font was presented centered on a white display with a variable duration ranging from 750 ms to 1,250 ms. Each letter was preceded by a fixation cross that was presented for 250–750 ms. Durations for fixation and stimulus presentation added up to a maximum of 1,500 ms. Participants were instructed to respond to any letter except *X* by pressing the ↓ button as fast as possible (go trial), but not to respond when the letter *X* occurred (no-go trial). Participants carried out the go/no-go task for 10 min test time, and thus completed at least 400 trials with a 20% rate of no-go stimuli. Note that the number of trials for each participant depended on the participant's individual response time. The go/no-go task described above was used in pretest, posttest, and follow-up testing alike. For the training sessions, a 10-min version of the task was used. In the first training session, stimulus durations and no-go stimulus rates were adjusted based on pretesting performance values of each individual (stimulus duration: mean reaction time [RT] + 250 ms; no-go stimulus rate in percent: 0.20 ± factor reflecting the relation of correct rejections and false alarms; algorithm available upon request). Extreme values were adjusted, so that stimuli were presented for at least 350 ms and a maximum of 1,000 ms, and the rate of no-go stimuli was more than 5%, but less than 35%. For the nonadaptive group, these parameters were used as fixed values for all training sessions; that is, task demand did not change and thus individuals could not reach their capacity level. In the adaptive training group, however, these parameters served as starting values for Training Session 1 and were then adjusted on a trial-by-trial basis. Performance was saved at the end of each training session to provide a starting value for the next session. To provide an individual adaptation for the given task, stimulus duration increased by 50 ms, if the participant made an error, and decreased by 50 ms in case of correct responses to no-go stimuli. Stimulus duration was also adjusted according to the participant's RT in go trials; that is, the duration for the next stimulus decreased by 50 ms, if the participant's response was faster than the average reaction in the session's previous trials, and increased by 50 ms, if the response was slower. A lower threshold of 250 ms stimulus duration was determined to guarantee stimulus detection. Additionally, the no-go stimulus rate increased by 5%, if the subject wrongly responded to a no-go stimulus, and decreased by 5%, if a no-go stimulus was correctly ignored. No-go stimulus rates in-

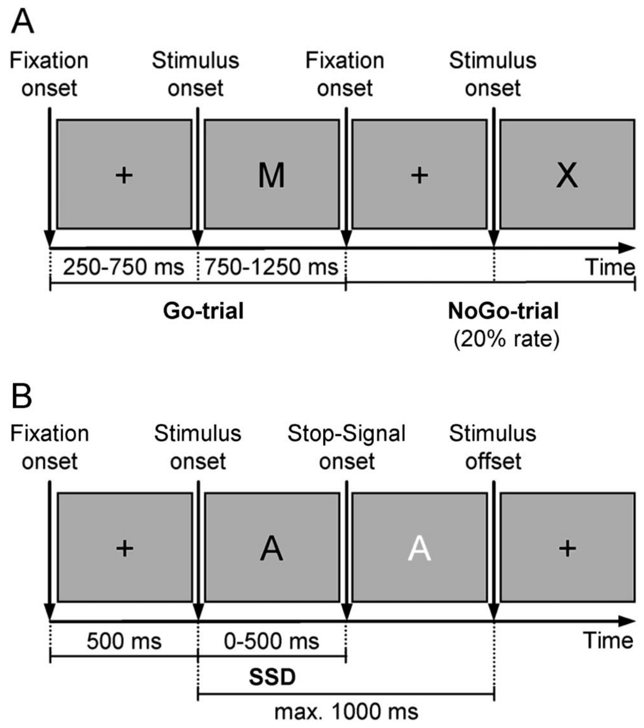

Figure 1. Inhibitory control training tasks (go/no-go [A] and stop signal [B]) in the pretest, posttest, and follow-up test session. For the training sessions, an adjusted version of the tasks was used that varied between adaptive and nonadaptive training groups: (A) Initial stimulus presentation time and no-go trial rate were adjusted according to individual's pretesting performance. Further, for the adaptive training group, both values were continuously adjusted on a trial-by-trial basis according to the individual's training performance: When an incorrect no-go response occurred, no-go trial rate was raised by 5% and stimulus presentation time was lengthened by 50 ms, while for correct no-go responses, no-go trial rate was reduced by 5% and stimulus presentation time was shortened by 50 ms. Furthermore, stimulus presentation time decreased by 50 ms if the participant's response was faster than the average reaction in the session's previous trials and increased by 50 ms if the response was slower. (B) Initial stop-signal delay (SSD; varied from 0 to 500 ms in 100-ms steps) was adjusted according to individual's pretesting performance. Further, SSD of the adaptive training group was continuously adjusted on a trial-by-trial basis according to the individual's training performance: When an incorrect stop response occurred, SSD was shortened by 50 ms; for correct stop responses SSD was lengthened by 50 ms. Note that the white letter (stop signal) actually appeared in red. max. = maximum.

creasing to more than 50% or decreasing to less than 2.5% were corrected to meet these thresholds, respectively.

Stop-signal task (Training Task 2). In the stop-signal task (see Figure 1B), a series of black capital letters in 72-point Arial font was presented on a white screen. Participants were instructed to discriminate between vowels and consonants by pressing the buttons ← (for vowel) and → (for consonant). Letters were presented centered on the screen for up to 1,000 ms and were preceded by a fixation cross that was presented for 500 ms. The stimulus set comprised vowels and consonants in equal number, but the letter *X* was excluded from the stimulus set because it served as a no-go stimulus in the go/no-go task. In addition, participants were instructed to suppress their response to the stimulus if the letter appeared in red font color or changed its color from the default black to red (stop signal) during the trial (stop trial). In stop-trials the stop-signal delay (SSD), that is, the delay between stimulus and stop signal appearance, varied from 0 ms to 500 ms in 100-ms increments. In pretest, posttest, and follow-up testing, the participants performed the stop-signal task for 10 min test time, during which at least 400 trials were completed. The rate of stop trials was about 25%. The adjusted stop-signal task used in the training sessions, which also lasted for 10 min, had identical stimulus and fixation durations but different rates of SSD. For participants of the nonadaptive group, the individual SSD was adjusted according to pretesting performance (based on mean RT and the relation between correct rejections and false alarms; algorithm available upon request), and this value was constantly used during the nine training sessions. In contrast, for the adaptive training group, this starting SSD was subsequently adjusted on a trial-by-trial basis, and performance indicators were saved at the end of each training session to serve as starting parameters in the following session. To adjust task difficulty continuously, the SSD decreased by 50 ms, if participants erroneously executed the discrimination task during a stop trial. If the response was correctly inhibited, however, the SSD increased by 50 ms.

Stroop task (near-transfer). Participants performed a classic Stroop task during pretest, posttest, and follow-up testing for 5 min (200 trials minimum), respectively. Different color names ("GREEN," "RED," "BLUE") and a neutral stimulus "++++" were successively presented centered on the screen in varying font colors (red, green, or blue) in 72-point Arial font for up to 1,000 ms on a white screen. Stimuli were preceded by a fixation cross for 500 ms. Participants were instructed to identify the font color of the presented stimuli by pressing the arrow keys ← for red fonts, ↓ for green fonts, and → for blue fonts. Three types of trials were conducted: (a) in congruent trials, font color and word meaning matched (e.g., "GREEN" written in green fonts); (b) in incongruent trials, font color and word meaning mismatched (e.g., "RED" written in blue); and (c) in neutral trials, the stimulus "++++" was presented in one of the font colors as outlined above. Training of IC should particularly affect performance in response to incongruent trials where response conflict is high.

Fluid intelligence (far-transfer). To measure Gf, two test halves of Raven's APM (Raven, 1990a; Raven, 1990b) were conducted and used for pre- and posttesting, respectively. APM item pairs of comparable difficulty were selected, and the two items of each pair were randomly assigned to one of two test halves, resulting in two APM test halves of 18 items, each with a mean item difficulty of .77. After participants finished the 5-min lasting practice module (Set 1) of the APM, they had 20 min to perform the test items. This is considered a valid measure of APM performance (Hamel & Schmittmann, 2006). In the follow-up testing, to prevent retest effects of the APM, the comparable matrices scale of the reasoning module of the well-established Intelligenz-Struktur-Test 2000 R Version C (IST 2000–R; Liepmann, Beauducel, Brocke, & Amthauer, 2007) was conducted. After a short practice block, participants had 10 min to complete the 20 items of this scale. Additionally, given the time between training and follow-up testing suggesting at best small transfer effects, the Wiener Matrizen-Test (WMT; Formann & Piswanger, 1979) was used to proof whether possible transfer can be replicated on a similar Gf test. The WMT is an adapted version of the Raven progressive matrices that conforms to the Rasch model (Rasch, 1960). It contains

24 matrices with increasing task difficulty and was administered with a time limit of 15 min.

Positive and negative affect scales. Because possible differences in positive and negative affective states can severely impact on task engagement and performance, the participant's current mood states were rated on the German version of the Positive and Negative Affect Schedule State version (PANAS–S; Krohne, Egloff, Kohlmann, & Tausch, 1996) at the beginning of each test and training session. The PANAS–S is sensitive to mood fluctuations and, according to Watson, Clark, and Tellegen (1988), can provide information of whether the experimental groups differ with regard to the extent they feel enthusiastic, active, and alert (positive affect), with higher scores reflecting states of energy, full concentration, and pleasurable engagement; or the extent they feel distressed, afraid, and nervous (negative affect), reflecting unpleasurable engagement.

Statistical Analyses

In a first step, it was examined whether the training procedure and the parameter adjustments in the IC tasks were suitable to train IC. For both training tasks, repeated-measures analyses of variance (ANOVAs) were conducted with the nine training sessions as repeated factors. Specifically, for the go/no-go task, it was determined whether the two parameters adjusted in the adaptive training group in a trial-by trial fashion (i.e., stimulus duration and no-go stimulus rate) decreased over the nine test sessions, which would indicate that individuals could handle increasing task demands. Similarly, for the stop-signal task, it was examined whether the adjustment of task difficulty by SSD increase and decrease, depending on their task performance, was suitable to train individuals of the adaptive training group in their ability to accurately inhibit a go response after a stop-signal occurred.

In a second step, mixed-model ANOVAs were conducted to examine (a) whether individuals who attended a 3-week ongoing adaptive training of the go/no-go and stop-signal IC tasks showed training-induced task improvements (as indicated by mean RT on go trials and by commission error rate, d', and stop-signal RT as more specific indicators of IC) compared to the nonadaptive (active control) group and, to provide comparison with other studies, to the passive control group; and (b) whether they differed with regard to their performance in a nontrained Stroop task (near-transfer) and psychometric Gf (far-transfer).

All analyses were done with SPSS 18.0 and are described in more detail in the Results section. Greenhouse–Geisser corrected degrees of freedom were applied where appropriate, and original degrees of freedom, epsilon adjustment values, and corresponding F values were reported. Note that individuals who exhibited overall error rates of more than 40% or RTs larger than 1,000 ms in the training/near-transfer tasks were discarded from the respective analyses. With our sample size of at least 117 for the comparison of pre- with posttesting and of at least 107 for the comparison of pretest with follow-up testing, we were able to detect a minimum effect size of about $f = .13$ and $f = .14$, respectively, with an alpha of .05 and a power of .80. Thus, relatively small effects could be detected.

Results

Descriptives

Descriptive statistics are shown in Table 1. Means and standard deviations are reported for the two training tasks (go/no-go, stop-signal), the near-transfer task (Stroop), and the Gf measures separated by the three groups (adaptive, nonadaptive, passive control) and the three time slots (pretest, posttest, follow-up). Moreover, intercorrelations of the used tasks at pretest are depicted in Table 2. The results indicate that the two training tasks as well as the training tasks and the Stroop task are significantly correlated with each other. However, none of the IC tasks was significantly

Table 1

Means (and Standard Deviation) of Performance in the Two Training Tasks (Go/No-Go, Stop Signal), the Near-Transfer Task (Stroop), and the Fluid Intelligence (Gf) Measure Separated by the Three Groups (Adaptive, Nonadaptive, Passive Control) and the Three Time Slots (Pretest, Posttest, Follow-Up Test)

	Go/no-go				Stop signal				Stroop			Gf			
Time slot	*N*	RT	*d′*	cE%	*N*	RT	cE%	SSRT	*N*	RT	SE	*N*	APM	IST	WMT
Adaptive															
Pretest	43	425 (47)	3.40 (0.63)	23 (12)	43	665 (92)	34 (18)	351 (61)	43	625 (103)	70 (50)	43	13.9 (1.8)		
Posttest	43	359 (39)	3.40 (0.64)	29 (17)	43	538 (97)	50 (17)	341 (47)	43	576 (95)	49 (39)	43	14.9 (2.4)		
Follow-up	41	353 (33)	3.34 (0.62)	34 (18)	41	539 (82)	51 (16)	334 (45)	41	562 (83)	54 (46)	41		12.8 (2.7)	19.0 (2.9)
Nonadaptive															
Pretest	38	419 (37)	3.27 (0.56)	24 (12)	37	656 (72)	35 (15)	358 (76)	37	628 (82)	64 (48)	38	13.1 (2.1)		
Posttest	38	349 (34)	3.20 (0.55)	34 (16)	37	526 (83)	52 (16)	342 (53)	37	554 (58)	36 (22)	38	14.3 (2.3)		
Follow-up	32	340 (28)	3.23 (0.60)	38 (19)	31	532 (73)	51 (17)	342 (62)	33	545 (63)	36 (23)	34		12.1 (2.1)	17.7 (3.1)
Passive control															
Pretest	38	414 (28)	3.26 (0.54)	25 (13)	37	647 (84)	39 (18)	366 (65)	37	622 (89)	54 (40)	39	13.2 (1.9)		
Posttest	38	380 (34)	3.27 (0.67)	30 (15)	37	580 (96)	48 (21)	346 (67)	37	565 (63)	44 (38)	39	14.5 (1.9)		
Follow-up	35	366 (26)	3.25 (0.60)	33 (15)	36	578 (85)	49 (17)	366 (54)	36	555 (63)	36 (45)	37		12.7 (2.5)	18.2 (3.6)

Note. Descriptive statistics include the respective number of observations that were included in the subsequent statistical analyses. RT = reaction time (in milliseconds) on correct trials; cE% = error rate (in percent) on no-go/stop trials; SSRT = stop-signal reaction time; SE = stroop effect; APM = Raven's Advanced Progressive Matrices; IST = Intelligenz-Struktur-Test; WMT = Wiener Matrizen-Test.

Table 2
Intercorrelations of Training and Transfer Measures at Pretest (Time 1)

Variable	1	2	3	4	5	6	7	8	9
1. GNG_RT	—								
2. GNG_*d′*	$.30^{**}$	—							
3. GNG_cE%	$-.56^{***}$	$-.78^{***}$	—						
4. SST_RT	$.47^{***}$	$.45^{***}$	$-.54^{***}$	—					
5. SST_cE%	$-.39^{***}$	$-.55^{***}$	$.61^{***}$	$-.93^{***}$	—				
6. SST_SSRT	−.01	$-.31^{***}$	$.28^{**}$	$-.32^{**}$	$.50^{***}$	—			
7. Stroop_RT	$.47^{**}$	.08	$-.22^{*}$	$.24^{**}$	$-.18^{*}$	.09	—		
8. Stroop effect	$.24^{**}$	−.10	−.01	−.08	.11	$.15^{\dagger}$	$.51^{***}$	—	
9. Gf	−.05	−.03	.07	−.10	.06	.05	−.12	−.01	—

Note. GNG = go/no-go task; RT = mean reaction time; cE% = error rate (in percent) on no-go/stop trials; SST = stop-signal task; SSRT = stop-signal reaction time; Gf = fluid intelligence.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

associated with the Gf measures, suggesting far-transfer effects to be less likely.

Between-Group Differences of Possible Confounding Factors

It was tested whether training and control groups differed with regard to the potentially confounding factors sex, age, sleep duration, smoking status, malaise, and caffeine and alcohol consumption, as well as positive and negative affective mood states. Mixed-model ANOVAs revealed that the adaptive training group did not differ from either the nonadaptive training group or the passive control group across the sessions (all $p > .10$).

Effectiveness of the Training Procedure

In a next step, we examined whether the training procedure was suitable to train IC in the adaptive training group. For both training tasks, repeated-measures ANOVAs were conducted with the nine training sessions as repeated factors. As outlined above, adaptive training of IC in the go/no-go task was operationalized by the adjustment of stimulus duration and no-go stimulus rate. Indeed, in the adaptive training group, stimulus duration at the beginning of the training sessions significantly decreased from 673 ms in Training Session 1 to 576 ms in Session 9 ($p < .001$; see also Figure 2, first panel). In view of the adjustment criteria (depending on the participant's RT in go trials and correct responses to no-go trials), the decrease of stimulus duration was mainly due to a substantial decrease in mean RT observed for the adaptive training group (and also for the nonadaptive training group) over the nine test sessions (see Figure 2, third panel). Further, a significant decrease in the no-go rate over the nine test sessions was observed when Session 1 was compared to the mean value of Sessions 2–9 ($p = .046$) and a trend when Session 1 was compared to Sessions 4–9 where performance had leveled out ($p = .079$; see Figure 2, second panel). However, with respect to both parameters, the adaptive training group significantly differed from the fixed values of the nonadaptive group (stimulus duration: $p < .05$ for Training Sessions 2–9; no-go rate: $p < .05$ for Training Sessions 2, 3, 6, and 8), indicating that task difficulty was much higher during the training sessions for the adaptive compared to the nonadaptive group.

In the stop-signal task, adaptive training of IC was obtained by the adjustment of the SSD that increased when the response was correctly inhibited during a stop trial. According to the horse race model (Logan & Cowan, 1984), at longer delays, it becomes more and more difficult to cancel the go process as the stop process starts later, and thus the go process finishes when the stop process is still going on. Effective training should be reflected in the stop-signal RT (SSRT), reflecting the speed of the inhibitory process after the stop-signal arose. As the adaptive adjustment produced response probabilities of about 50%, SSRT could be calculated by subtracting the mean SSD from the mean go RT (see Verbruggen & Logan, 2009), with lower values indicating higher IC. Against our expectations, individuals of the adaptive group showed a significant and relatively continuous increase of SSRT from 164 ms at Training Session 1 to 213 ms at Training Session 9 ($p < .001$; see Figure 2, last panel). This increase, however, was mainly due to a large and highly significant ($p < .001$) decrease in mean RT from Session 1 (597 ms) to Session 9 (508 ms; see Figure 2, fifth panel), which occurred at the cost of an increase in mean SSD over the nine test sessions (see Figure 2, fourth panel). In view of Figure 2, individuals of the adaptive group showed the expected increase in mean SSD in Training Sessions 1 and 2, which was reflected in low SSRT. Over the subsequent trainings sessions, however, SSD decreased, accompanied by a strong decrease in RT and thus an increase in SSRT, too. Thus, one might argue that their focus had changed from slower but more accurate toward faster but less accurate responding.

Moreover, it was examined whether the negative affective mood states substantially changed over the nine test sessions. As outlined above, the Negative Affect scale is a dimension of distress and a lack of engagement that might severely compromise training gains in the adaptive training group. However, the analysis neither gives evidence for an increase in negative affect nor suggests any differences between the adaptive and nonadaptive groups (all $p > .05$).

Between-Group Differences in the Training Tasks

Next, we examined whether the adaptive training group showed larger performance increase at the posttest (Time 2 [T2]) and the follow-up test session (Time 3 [T3]) than the nonadaptive and passive control groups. No group-specific differences should occur at Time 1 (T1). Mixed-model ANOVAs were conducted for the following indicators of overall performance in the go/no-go and stop-signal tasks: Commission errors (cE%) indicating failures to

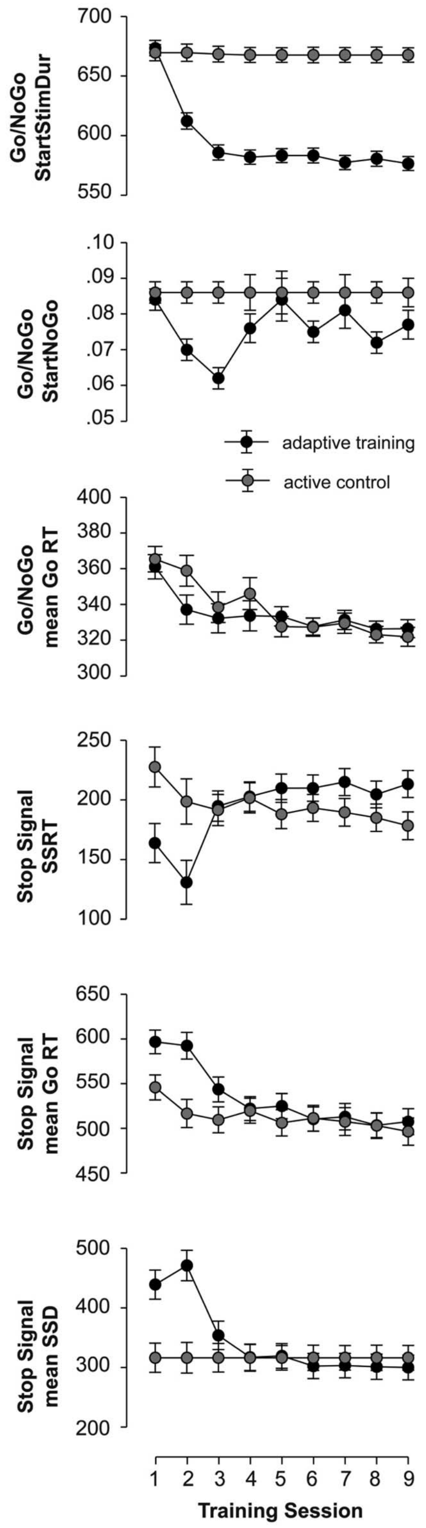

Figure 2. Start adjustment parameters of the adaptive training and group-related go/no-go and stop-signal task performance at training Sessions 1–9. StartStimDur = starting value of stimulus duration; StartNoGo = starting value of no-go trial rate; RT = reaction time; SSRT = stop-signal reaction time; SSD = stop-signal delay. Error bars represent standard error of the mean.

suppress the response in the no-go/stop trials and d' in the go/no-go task and SSRT in the stop-signal task were considered as specific indicators of response inhibition. Moreover, mean RT on correct go trials (RT) was considered, as analyses of the training sessions indicated that performance benefits were especially obtained with respect to RT. The ANOVAs included the within-subjects factors Time (T1 vs. T2/T3) and the between-subjects factor Group (adaptive training, nonadaptive active control, passive control). Note that the three test sessions were analyzed in two ANOVA models (Model 1: T1 vs. T2; Model 2: T1 vs. T3) to provide maximal sample size and power.

With respect to the go/no-go task performance, for the comparison of T1 with T2, there were no significant group differences for cE% and d' (all $p > .05$; for exact p values and effect sizes, see Appendix; see also Figure 3A, middle and right panels) indicating that adaptive training of the inhibition tasks did not cause substantial variance in these measures at the posttesting. When, however, comparing T1 with T3, the analysis revealed a significant Time × Group effect for cE% ($F_{2,\ 105} = 3.27$, $p = .042$, $\eta^2_p = .06$), which was due to a somewhat steeper increase of cE% at T3 compared to T1 for the nonadaptive group (see Figure 3A, right panel). As indicated by parameter estimates, these group differences, however, did not reach statistical significance (for all simple comparisons, $p > .05$). For mean RT on go trials in the T1–T2 comparison, the analysis showed a very large and highly significant Time × Group interaction effect ($F_{2,\ 116} = 14.77$, $p < .001$, $\eta^2_p = .20$). As expected, no group differences occurred at T1 (adaptive vs. nonadaptive: $p = .515$; adaptive vs. passive control: $p = .204$). At T2, the adaptive training group showed large RT decrease (T1 vs. T2: $F_{1,\ 116} = 175.56$, $p < .001$, $\eta^2_p = .60$). However, this decrease was similarly high for the nonadaptive active control group (T1 vs. T2: $F_{1,\ 116} = 174.17$, $p < .001$, $\eta^2_p = .60$; adaptive vs. nonadaptive: $p = .235$). Significant group differences also occurred when the adaptive training group was compared to the passive control group that exhibited a much smaller RT decrease ($F_{1,\ 116} = 39.05$, $p < .001$, $\eta^2_p = .25$; adaptive vs. passive: $p = .007$), which clearly indicates that practicing the tasks impacts RT performance at posttesting. When comparing pretest (T1) and follow-up test session (T3), the observed group differences in RT decrease remained stable, even 4 months after the training ($F_{2,\ 105} = 10.48$, $p < .001$, $\eta^2_p = .17$). The group pattern at T3 went parallel to that observed at T2 (adaptive vs. nonadaptive: $p = .076$; adaptive vs. passive control: $p = .049$). The group-related RT performance in the go/no-go task at T1, T2, and T3 is depicted in Figure 3A (left panel).

Concerning group-related performance in the stop-signal task, similar effects were observed (see Figure 3B). For the comparison of T1 and T2, there was a large Time × Group interaction effect for mean RT on go trials ($F_{2,\ 114} = 10.46$, $p < .001$, $\eta^2_p = .16$). Again, no group differences were observed at T1 (adaptive vs. nonadaptive: $p = .656$; adaptive vs. passive control: $p = .358$). As in the go/no-go task, larger RT improvements from T1 to T2 were observed for the adaptive group (T1 vs. T2: $F_{1,\ 114} = 151.78$, $p < .001$, $\eta^2_p = .57$) compared to the passive control group (adaptive vs. passive: $p = .045$). However, the nonadaptive group showed similar RT decrease (T1 vs. T2: $F_{1,\ 114} = 139.39$, $p < .001$, $\eta^2_p = .55$), and thus again, the adaptive and nonadaptive groups did not differ from each other (adaptive vs. nonadaptive: $p = .544$). Notably, the group-related differences in RT decrease remained

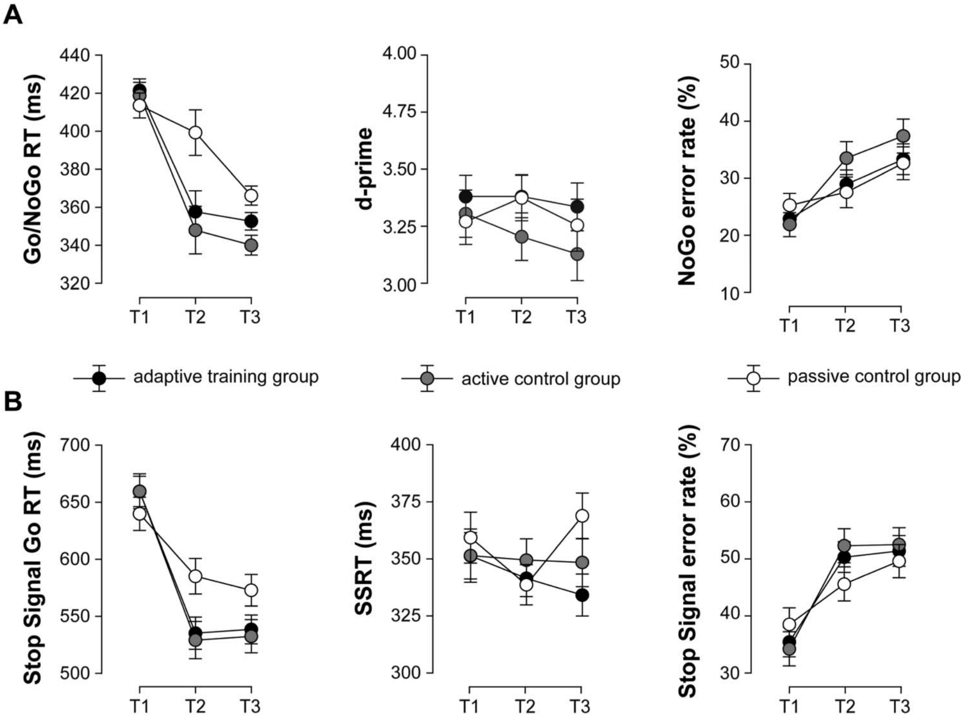

Figure 3. Group-related training task performance at pretest (Time 1), posttest (Time 2), and follow-up testing (Time 3) for the go/no-go task (A) and the stop-signal task (B). The three test sessions were analyzed in two analysis of variance models (Model 1: Time 1 vs. Time 2; Model 2: Time 1 vs. Time 3) to provide maximal sample size and power. RT = mean reaction time; SSRT = stop-signal reaction time; T1–T3 = Time 1–Time 3. Error bars represent standard error of the mean.

stable even 4 months after training ($F_{2,\ 105} = 7.36$, $p = .001$, $\eta^2_p = .12$), as suggested by a larger RT decrease from T1 to T3 for the adaptive compared to the passive control group ($p = .034$). No differences between adaptive and nonadaptive groups were observed ($p = .733$). Moreover, similar to the result pattern obtained for the go/no-go task, no significant group effects occurred for SSRT, providing a more specific indicator of response inhibition (all $p > .05$; for exact *p* values and effect sizes, see Appendix; see also Figure 3B, middle panel). With respect to cE%, a significant Time × Group effect occurred for the comparison of T1 and T2 ($F_{2,\ 114} = 3.51$, $p = .033$, $\eta^2_p = .06$). As in the go/no-go task, however, parameter estimates indicated that there were no significant group differences at either T1 or T2. All groups showed larger cE% rates at the posttest than at the pretest, but the slope of the adaptive group and, even more pronounced, that of the nonadaptive group was somewhat steeper than that of the passive control group (see Figure 3B, right panel).

Between-Group Differences in Nontrained Tasks

Near-transfer. To examine transfer on Stroop performance, we performed mixed-model ANOVAs for overall Stroop RT as well as the Stroop congruency effect (incongruent minus congruent RT). The ANOVAs included the within-subjects factors Time (T1 vs. T2/T3) and the between-subjects factor Group (adaptive, nonadaptive, passive control). Comparing pre- and posttesting (T1 vs. T2) and pretest with follow-up testing (T1 vs. T3), no significant group-related main or interaction effect emerged (all $p > .05$; for exact *p* values and effect sizes, see Appendix), suggesting that the adaptive training group showed no substantial differences with regard to either the active control or the passive control group in Stroop performance. The group-related effects on Stroop task performance at T1, T2, and T3 are depicted in Figure 4.

Far-transfer. Possible training effects on Gf were explored by a mixed-model ANOVA with Time (T1 vs. T2/T3) as within-subjects factor. The comparison of T1 with T3 was conducted for the averaged score of the two Gf measures used in the follow-up test session (IST 2000–R, WMT). All Gf scores were *z*-transformed to ensure comparability. None of the ANOVA models revealed significant Time × Group interactions or group main effects (all $p > .05$; see Appendix).

Discussion

IC is considered a key factor of goal-directed behavior (Barkley, 1997; Hasher & Zacks, 1988; Ridderinkhof et al., 2004), and IC deficits have been shown to play a crucial role in neuropsychiatric syndromes such as ADHD (Barkley, 1997), compulsive-affective disorder (Bannon et al., 2002), and addiction (Bechara, 2005) and in the decline of prospective memory across the life span (Kliegel et al., 2008).

Thus, it would be desirable to know whether and to what extent IC can be improved by cognitive training. However, training studies addressing IC yielded inconsistent results concerning their

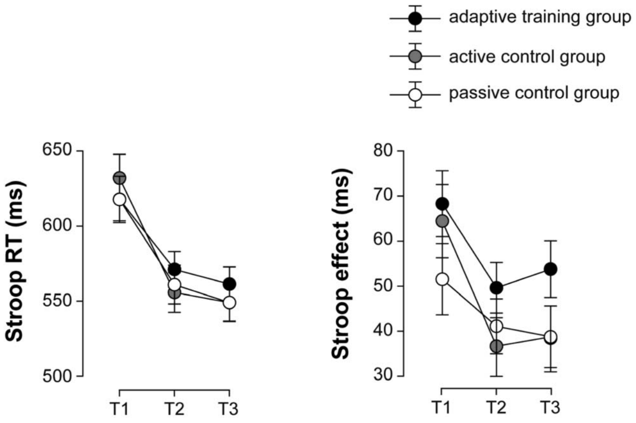

Figure 4. Group-related Stroop performance (near-transfer) at pretest (Time 1), posttest (Time 2), and follow-up testing (Time 3) measured by mean reaction time (RT) and the Stroop congruency effect (incongruent minus congruent RT). The three test sessions were analyzed in two analysis of variance models (Model 1: Time 1 vs. Time 2; Model 2: Time 1 vs. Time 3) to provide maximal sample size and power. T1–T3 = Time 1–Time 3. Error bars represent standard error of the mean.

feasibility to produce training and transfer effects (see Morrison & Chein, 2011; Shipstead et al., 2010; Thorell et al., 2009), focused on small samples with executive deficits (Holmes et al., 2009; Klingberg, 2010), frequently used passive control groups (see Shipstead et al., 2012), or did not directly train IC, as IC tasks were used to test transfer effects of WM training (see Karbach & Kray, 2009; Klingberg, 2010). Thus, the main goal of the present study was to provide an estimation of the effect size and stability of IC training in a healthy population. It was also of interest whether and to what extent the training caused near-transfer on untrained IC measures (Stroop). In addition, far-transfer on psychometric Gf was explored. Based on recent methodological advice (Klingberg, 2010; Moody, 2009; Shipstead et al., 2010, 2012; Sternberg, 2008), this study was designed as a randomized, double-blind pretest, posttest, and follow-up trial.

Between-Group Differences in the Training Tasks

The training of the two established IC tasks (go/no-go, stop-signal) was expected to result in training gains as indicated by overall task performance and more specific indicators of IC (e.g., error rate on no-go/stop trials, SSRT). When comparing performance gain from pre- to posttesting, the 3 weeks' training led to a large and highly significant decline in mean RT in both the adaptive and nonadaptive (active control) groups. The adaptive training group did not significantly differed from the active control group but showed a significant training gain with respect to the passive control group that even remained stable in the follow-up measurement 4 months later. This is similar to a number of previous studies on WM training that frequently observed training gains for the adaptive compared to the passive group but not to the active controls (see Klingberg, 2010; Morrison & Chein, 2011; Shipstead et al., 2012). Given that the present study used a young healthy sample of lower variance for IC measures relative to samples with executive deficits, this large effect would suggest a high effectiveness of the present training. However, as discussed for WM trainings, only differences between the adaptive training group and the active controls (nonadaptive group) would allow firm conclusions on the possible benefits of an adaptive IC training, as both groups were administered to the same treatment except that the adapted version of the tasks pushes individuals to their maximal level of task performance and thus is far more demanding than the nonadaptive one (see Klingberg, 2010; Shipstead et al., 2010, 2012).

In contrast, comparisons between the adaptive group and no-contact controls who are not involved in any similar task in between the pre- and posttest sessions can be somewhat problematic. For example, they are not as often confronted with the tasks, the laboratory setting, and the experimenters than the other groups, and they may also have recognized not to receive experimental treatment. As a consequence, effects between the training group and passive controls that occurred after the training can potentially (although not necessarily) be explained by alternative sources of variation that can influence task performance and may have led to the observed differences between the groups. Such unwanted variance that, for example, relate to a person's engagement in the study (e.g., Hawthorne/placebo or expectancy effects) can affect the internal validity of training studies, as discussed recently (Redick et al., 2013; Shipstead et al., 2010, 2012). Although the influence of several task-relevant confounds has been controlled, such as task-relevant mood fluctuations relating to the extent of (un)pleasurable engagement (Krohne et al., 1996; Watson et al., 1988), this cannot rule out that other concomitant factors may have contributed to the observed effect, as outlined above. Thus, given that the adaptive training group differed from the passive control group but not from the active control group in RT performance of the IC tasks at posttest and follow-up testing, training effects that can be conclusively attributed to real changes in IC function could not be observed. Alternatively, performance differences between the training group and passive control group may be attributed to practice effects independent of changes in IC or to group-related differences in task engagement or motivation.

This conclusion is also supported by the fact that training gains were only observed with respect to RT but not regarding the more specific parameters of IC such as error rate on no-go/stop trials, *d′* (go/no-go task), or SSRT (stop-signal task). This poses the question whether the training of IC tasks can actually result in training effects of IC.

The Effectiveness of Parameter Adjustments in the Adaptive Training

To evaluate whether the adaptive training procedure was appropriate to increase IC performance, it was examined whether the adaptive training group benefited from performance-dependent parameter adjustments over the course of training. Indeed, for the go/no-go task, the expected performance improvements were obtained during training. Specifically, this was indicated by a large decrease in stimulus duration and by a significant decrease in the no-go rate over the nine test sessions. However, in the stop-signal task, adjustment of the SSD as a function of task performance did not result in continuously increasing SSDs, as would have been expected. Consequently, SSRT reflecting the speed of the inhibitory process after a stop-signal occurs increased rather than decreased. A closer look on the training data shows that during

Training Sessions 1 and 2, individuals of the adaptive training group "trained" as expected. That is, they showed an increase in mean SSD, which was reflected in decreased SSRTs. In Training Sessions 3 and 4, however, SSD decreased (and remained invariant from Session 5 to Session 9), which was accompanied by a strong RT decrease and thus an increase in SSRT. In light of these data, one might argue that the response patterns of the adaptive group members changed from a slower but more accurate toward a faster but less accurate responding. Given that individuals can focus either on the go trials or on the stop trials or on both to meet the requirements of the task, it may be conceivable that with increasing training (and thus task demand) individuals have shifted their focus to those stimuli that can be most easily controlled (i.e., the go stimuli) in order to accomplish the task as well as possible. This could have been further promoted by the fact that the participants were instructed to be as fast and accurate as possible (as is common for these kind of tasks). Overall, this may have led to shorter response latencies in go trials at the cost of accuracy in the stop trials, indicating that speed–accuracy trade-offs may play a critical role here. This impression was also supported when stop-signal task improvements from pre- to posttest session were considered. Here overall RT in all groups (most pronounced in the adaptive and nonadaptive groups) decreased, whereas commission error rate (i.e., false alarms on no-go/stop trials) increased, which was similarly observed for the go/no-go task. Because of this, training gains in specific indicators of response inhibition might be more difficult to obtain than compared to the training of WM capacity, where task performance and load adjustments only depend on correct target responses. This might be particularly true for the stop-signal task, as we observed a strategy change during the course of training toward speeded responses on more easy-to-control go trials, as outlined above.

Near- and Far-Transfer as a Function of Training

As recently summarized (Klingberg, 2010; Morrison & Chein, 2011; Shipstead et al., 2012), several studies reported training increases to significantly explain variance in untrained behavioral measures. However, as outlined above, training effects in the present study only occurred when the adaptive training group was compared with passive controls and also only with respect to overall task performance as indicated by mean RT. That is, potential benefits of the adaptive training group in nontrained tasks (near- and far-transfer) cannot be readily attributed to changes in IC (see Redick et al., 2013; Shipstead et al., 2012). Nevertheless, to provide a full picture of the effects and comparability with previous studies using passive control groups, as recommended (Redick et al., 2013), we tested whether the adaptive training group differed from the two control groups with respect to the nontrained Stroop task (near-transfer) and to Gf (far-transfer). Moreover, larger transfer effects for the training compared to the active control group would suggest that their similar performance increase in the training tasks could be attributable to ceiling effects. Because performance gains in the training tasks from T1 to T2 were observed for overall RT but not for specific IC markers, a potential transfer effect on Stroop task was more likely for overall Stroop RT than for the Stroop congruency effect measuring IC more specifically. However, neither for posttest nor for the follow-up testing significant group differences in any of the Stroop performance indicators were detected.

Similarly, no transfer on psychometric Gf was observed. A few studies also reported successful far-transfers of executive training (mainly WM tasks) on Gf performance (for reviews and critique, see Klingberg, 2010; Morrison & Chein, 2011; Shipstead et al., 2010, 2012). It might be that the high conceptual analogy between WM and Gf (e.g., Friedman et al., 2006) favors WM as basis for training-induced far-transfer (Sternberg, 2008), although transfer on Gf has been recently reported using a task-switching training (Karbach & Kray, 2009). However, the potential of WM training to cause transfer due to changes in WM has been recently challenged (Chooi & Thompson, 2012; Melby-Lervag & Hulme, 2013; Redick et al., 2013; Shipstead et al., 2010, 2012). Moreover, it is obvious that the proof of far-transfer to a rather broad and distant concept such as Gf is undeniably more complicated than proving near-transfer to performance in conceptually closer tasks such as the Stroop (see also wide-strength dilemma; Perkins & Salomon, 1989).

Overall, although in the adaptive training group, training effects were observed relative to passive controls, this cannot be conclusively attributed to increases in the IC function, as there were no differences with regard to the active controls and as these differences were related to overall task performance but not to specific markers of IC functioning.

Should Null Effects of IC Training Be Generalized?

In contrast to our healthy sample, which showed lower variability in executive functions, IC training and transfer effects (differences between training group and active controls) may be possible for individuals with cognitive deficits such as children with ADHD and stroke patients, where a potential for plasticity and neurocognitive reorganization can be expected (see, e.g., Lovden, Backman, Lindenberger, Schaefer, & Schmiedek, 2010). For example, Klingberg et al. (2005) reported improvements for children with ADHD in both RT and accuracy of WM training on untrained Stroop performance of $d = 0.34$ for the posttesting and $d = 0.25$ for the follow-up testing 3 months after the training. However, Karbach and Kray (2009) found transfer on Stroop performance using a task-switching training in healthy samples of different ages and even on Gf, but observed no such transfer in ADHD children (Kray, Karbach, Haenig, & Freitag, 2012). Further, training and near-transfer may also occur in the elderly (e.g., inhibitory deficit theory Hasher & Zacks, 1988) with age-related deficits in executive control functions (Lovden et al., 2010; Zinke, Zeintl, Eschen, Herzog, & Kliegel, 2012). In this context, future research should also examine whether the effectiveness of trainings is moderated by factors that explain interindividual variability and plasticity in IC such as genetic variations (e.g., Enge, Fleischhauer, Lesch, Reif, & Strobel, 2011; Meyer-Lindenberg et al., 2006) or temperamental factors.

As a general recommendation, in addition to control for violations of the internal validity of executive function training by using appropriate designs, future studies may control for other important sources of unwanted variance, such as systematic method variance. Training tasks and (near-) transfer tasks are typically based on conceptually similar requirements (e.g., similar motor responses). This, in turn, may lead to an overestimation of the size of training and transfer effects. In order to separate method- from construct-specific variance, the use of multitrait–multimethod designs, in which the variables of interest were assessed by several different methods, would be promising (Campbell & Fiske, 1959). Thus, the function of interest (here IC) should be measured not only by computerized tasks but also by test inventories and questionnaires (e.g., Rothbart, Ahadi, & Evans, 2000), behavioral observations (e.g., Klingberg et al., 2002), and endophenotypes (Gottesman & Gould, 2003), such as structural MRI data, positron emission tomography, and brain chemistry parameters, which would reflect the impact of training in brain circuits corresponding to the trained function. Thus, multitrait–multimethod designs would enable the a posteriori adjustment of the group-specific variance by the variance proportion caused by the different assessment methods, providing a valid estimate of the actual effect size.

Limitations

One may argue that the nine training sessions have limited possible training and transfer effects in the present study (e.g., that small effects would not have been detected). Although other training studies used a similar number of training sessions, there are a variety of studies using a larger number of sessions. In this context, however, recent studies and meta-analytic data of WM trainings suggest that, for example, irrespective whether the training included eight or 20 sessions, no significant transfer on nontrained mental abilities tests and Gf measures could be detected (Chooi & Thompson, 2012; see also, Melby-Lervag & Hulme, 2013; Redick et al., 2013). It is also notable that we used a comparatively large sample size, compared to most other training studies, that enabled us to detect relatively small effects. In addition, data on the development of the adjusted task parameters over the nine test sessions demonstrate performance changes in the adaptive training group, particularly in the first two to four test sessions that then leveled out. Thus, one might argue that more than nine test sessions would not have resulted in a more effective training. As outlined above, this might result from the fact that the individuals focused on speed rather than on accuracy. Notably, however, even with regard to the large decrease of RT in both IC tasks at T2, no significant transfer on RT in Stoop measures could be found in the present study, which renders the possibility of true training and transfer effects on more specific markers of IC also less likely. Thus, it seems rather unlikely that the limited effects in the present study can mainly be attributed to our number of training sessions.

Nevertheless, future research should ensure that individuals are more focused on stop trials, for example, by instructing them to be as accurate as possible in response to stop trials at the expense of response speed on go stimuli (i.e., to provoke a more conservative responding). Related to this, a further possible recommendation refers to the training algorithm. One may speculate that instead of adjusting the SSD trial by trial, as is common in training studies, training of response inhibition in the stop-signal task might have been more effective if SSD were increased by 50 ms after individuals had correctly inhibited their response in, for example, 10 subsequent stop trials (i.e., in a stepwise fashion). Such an adjustment would probably have allowed individuals to get more familiar with increasing task demand by increasing SSD.

Another possible limitation may refer to the use of the Stroop as an adequate near-transfer task. Although we did not observe substantial differences between the adaptive training group and the active controls after training that in turn would be a prerequisite for later transfer, this issue might be interesting for future research. We decided to utilize the Stroop as a near-transfer measure for the following reasons: The Stroop has been related to the efficiency to suppress prepotent responses, such as those that are already initiated or under way but need to be inhibited to effectively accomplish the task, and thus shares variance with motor response inhibition tasks (as, for example, reviewed in Friedman & Miyake, 2004; MacLeod, 1991; Miyake et al., 2000; Salthouse et al., 2003). In the present study, this could indeed be demonstrated by correlations between performance indices of the go/no-go and stop-signal tasks with the Stroop congruency effect at pretesting. Thus, the present data support the notion that performance of training tasks and near-transfer task is significantly associated. Although these correlations were small, it has been recommended that transfer tasks should share features and processes of the training task, but should be different enough from the training task to avoid mere practice effects (see also Jaeggi et al. 2008). Assuming that performance in the training tasks can lead to real changes in the IC function that cannot be attributed to practice/learning or other effects (Redick et al., 2013; Shipstead et al., 2010, 2012), we deemed the Stroop to be appropriate for testing potential transfer effects of IC training. Nevertheless, compared to the response inhibition tasks used for training, beyond doubt, the Stroop comprises several other aspects of executive control and is therefore considered a more complex IC task (see Friedman & Miyake, 2004; MacLeod, 1991; Miyake et al., 2000; Simmonds et al., 2008). Besides being a measure of suppressing prepotent responses, Stroop relates to performance in suppressing irrelevant stimuli, stimulus-triggered reflexive attention shifts, and irrelevant cognitions or habits (see, e.g., Salthouse et al., 2003). As underlying mechanism, an intern response competition is discussed (MacLeod, 1991), which is reflected by RT differences on congruent and incongruent trials (Dempster & Corkill, 1999b). However, alternative tasks for near-transfer would have been possible, such as the Simon task, which may provide a close proximity to the training tasks used (Hommel, 2011). Considering that interference suppression (Eriksen flanker task) and response inhibition (go/no-go task) appear to share overlapping neural bases in adults (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002), are highly correlated, and seem to fall along a single factor (Friedman & Miyake, 2004), other selective attention tasks related to inhibition, such as the Eriksen flanker or spatial Stroop task (Diamond, 2013), might appear as potential near-transfer measures for response inhibition training.

Furthermore, Stroop performance has been frequently found to explain substantial variance in several Gf measures such as reflected in moderate negative correlations with the Raven's matrices test (see Dempster & Corkill, 1999b; Kane & Engle, 2002). In contrast, however, in the present study, no correlations between the Stroop and Gf could be detected. Overall, for its shared variances with the trained tasks and based upon the conceptual proximity of Stroop to the IC training tasks, we regarded the Stroop as a possible near-transfer task that might also be seen suitable as a "connection" to more general cognitive performance.

Conclusion

In this work, the effectiveness of IC training in a healthy population was examined by comparing an adaptive training group with nonadaptive active controls. The study followed a randomized, double-blind pretest, posttest, and follow-up procedure and provided a relatively large sample size. For the sake of comparability with previous studies, a no-contact control group was also implemented. Although the adaptive group could substantially improve overall IC task performance after training, particularly with regard to RT, no differences to the active control group occurred, either at posttest or at follow-up testing. In contrast, as in a number of previous studies, significant training gain was found when the training group was compared to the passive control group, which, however, does not sufficiently control for possible confounds. Thus, contributing to the current debate, no conclusive evidence was found that this performance increase mirrors a true increase in IC function. However, the fact that the training gain was mainly related to overall RT performance may indicate that individuals appeared to be more concentrated on performance gains in the prepotent go trials but less to the stop trials to meet the requirements of the tasks as well as possible. Such a response strategy may be related to the typical design of the training tasks used but seemed to be particularly evident for the stop-signal task. Overall, this challenges the effectiveness of IC training on specific measures of response inhibition relative to the training of WM, as outlined above. By changes in task instructions and training algorithms, we provide possible recommendations that may provoke individuals to be more focused on stop trial performance instead of overall performance or response speed on go stimuli in order to obtain a more effective training of specific response inhibition indices (e.g., stop trial accuracy). However, even in light of the large RT decrease in the training tasks that was found at T2, neither significant differences between the training groups nor transfer effects could be detected, which challenge the expectation of true training-related changes for more specific response inhibition markers. Irrespective of this, the results do not rule out the possibility of training-related improvements in populations with IC deficits, particularly in those with the potential for neural plasticity and reorganization. This may also be true for factors that relate to interindividual variability and plasticity in IC function and thus may moderate the effectiveness of training, such as genetic variations and temperamental factors.

References

Anderson, M. C., & Weaver, C. (2009). Inhibitory control over action and memory. In L. Squire (Ed.), *Encyclopedia of neuroscience* (Vol. 5, pp. 153–163). Amsterdam, the Netherlands: Elsevier Academic Press. doi:10.1016/B978-008045046-9.00421-6

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8,* 170–177. doi:10.1016/j.tics.2004.02.010

Bannon, S., Gonsalvez, C. J., Croft, R. J., & Boyce, P. M. (2002). Response inhibition deficits in obsessive-compulsive disorder. *Psychiatry Research, 110,* 165–174. doi:10.1016/S0165-1781(02)00104-X

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94. doi:10.1037/0033-2909.121.1.65

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience, 8,* 1458–1463. doi:10.1038/nn1584

Borella, E., Carretti, B., Zanoni, G., Zavagnin, M., & De Beni, R. (2013). Working memory training in old age: An examination of transfer and maintenance effects. *Archives of Clinical Neuropsychology, 28,* 331–347. doi:10.1093/arclin/act020

Buehner, M., Krumm, S., & Pick, M. (2005). Reasoning = working memory ≠ attention. *Intelligence, 33,* 251–272. doi:10.1016/j.intell.2005.01.002

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron, 33,* 301–311. doi:10.1016/S0896-6273(01)00583-9

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–105. doi:10.1037/h0046016

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354–380. doi:10.1037/0033-2909.132.3.354

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17,* 193–199. doi:10.3758/PBR.17.2.193

Chooi, W.-T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence, 40,* 531–542. doi:10.1016/j.intell.2012.07.004

Ciesielski, K. T., & Harris, R. J. (1997). Factors related to performance failure on executive tasks in autism. *Child Neuropsychology, 3,* 1–12. doi:10.1080/09297049708401364

Clark, L., Cools, R., & Robbins, T. W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition, 55,* 41–53. doi:10.1016/S0278-2626(03)00284-7

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence, 32,* 277–296. doi:10.1016/j.intell.2003.12.002

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *Quarterly Journal of Experimental Psychology: Section A: Human Experimental Psychology, 51,* 305–320. doi:10.1080/713755761

Dempster, F. N., & Corkill, A. J. (1999a). Individual differences in susceptibility to interference and general cognitive ability. *Acta Psychologica, 101,* 395–416. doi:10.1016/S0001-6918(99)00013-X

Dempster, F. N., & Corkill, A. J. (1999b). Interference and inhibition in cognition and behavior: Unifying themes for educational psychology. *Educational Psychology Review, 11,* 1–88. doi:10.1023/A:1021992632168

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

Enge, S., Fleischhauer, M., Lesch, K. P., Reif, A., & Strobel, A. (2011). Serotonergic modulation in executive functioning: Linking genetic variations to working memory performance. *Neuropsychologia, 49,* 3776–3785. doi:10.1016/j.neuropsychologia.2011.09.038

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309–331. doi:10.1037/0096-3445.128.3.309

Enright, S. J., & Beech, A. R. (1993). Further evidence of reduced cognitive inhibition in obsessive-compulsive disorder. *Personality and Individual Differences, 14,* 387–395. doi:10.1016/0191-8869(93)90307-O

Enriquez-Geppert, S., Huster, R. J., & Herrmann, C. S. (2013). Boosting brain functions: Improving executive functions with behavioral training, neurostimulation, and neurofeedback. *International Journal of Psychophysiology, 88,* 1–16. doi:10.1016/j.ijpsycho.2013.02.001

Fisk, J., & Sharp, C. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26,* 874–890. doi:10.1080/13803390490510680

Formann, A. K., & Piswanger, K. (1979). *Wiener Matrizen-Test: Ein Rasch-skalierter sprachfreier Intelligenztest* [Viennese Matrices Test: A Rasch-scaled culture-fair intelligence test]. Weinheim, Germany: Beltz Test.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133,* 101–135. doi:10.1037/0096-3445.133.1.101

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17,* 172–179. doi:10.1111/j.1467-9280.2006.01681.x

Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 8301–8306. doi:10.1073/pnas.96.14.8301

Giancola, P. R., & Tarter, R. E. (1999). Executive cognitive functioning and risk for substance abuse. *Psychological Science, 10,* 203–205. doi:10.1111/1467-9280.00135

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 160,* 636–645. doi:10.1176/appi.ajp.160.4.636

Hamel, R., & Schmittmann, V. D. (2006). The 20-minute version as a predictor of the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement, 66,* 1039–1046. doi:10.1177/0013164406288169

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(08)60041-9

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12,* F9–F15. doi:10.1111/j.1467-7687.2009.00848.x

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology, 24,* 827–836. doi:10.1002/acp.1589

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica, 136,* 189–202. doi:10.1016/j.actpsy.2010.04.011

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 6829–6833. doi:10.1073/pnas.0801268105

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between *n*-back performance and matrix reasoning—Implications for training and transfer. *Intelligence, 38,* 625–635 doi:10.1016/j.intell.2010.09.001

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9,* 637–671. doi:10.3758/BF03196323

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12,* 978–990. doi:10.1111/j.1467-7687.2009.00846.x

Kliegel, M., Mackinlay, R., & Jager, T. (2008). Complex prospective memory: Development across the lifespan and the role of task interruption. *Developmental Psychology, 44,* 612–617. doi:10.1037/0012-1649.44.2.612

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14,* 317–324. doi:10.1016/j.tics.2010.05.002

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., . . . 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,* 177–186. doi:10.1097/00004583-200502000-00010

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24,* 781–791. doi:10.1076/jcen.24.6.781.8395

Kray, J., Karbach, J., Haenig, S., & Freitag, C. (2011). Can task-switching training enhance executive control functioning in children with attention deficit/hyperactivity disorder? *Frontiers in Human Neuroscience, 5,* 180. doi:10.3389/fnhum.2011.00180

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der “Positive and Negative Affect Schedule” (PANAS) [Investigations with a German version of the “Positive and Negative Affect Schedule” (PANAS)]. *Diagnostica, 42,* 139–156.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14,* 389–433. doi:10.1016/S0160-2896(05)80012-1

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6,* 299–305. doi:10.1016/S1364-6613(02)01923-X

Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). Manual (2. erweiterte und überarbeitete Aufl.)* [Intelligence Structure Test 2000–R (IST 2000–R). Manual (2nd expanded and revised ed.)]. Göttingen, Germany: Hogrefe.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users’ guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327.

Lovden, M., Backman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin, 136,* 659–676. doi:10.1037/a00200802010-12718-009

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203. doi:10.1037/0033-2909.109.2.163

Melby-Lervag, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49,* 270–291. doi:10.1037/a00282282012-12954-001

Meyer-Lindenberg, A., Buckholtz, J. W., Kolachana, B., Hariri, A. R., Pezawas, L., Blasi, G., . . . Weinberger, D. R. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 6269–6274. doi:10.1073/pnas.0511311103

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,* 49–100. doi:10.1006/cogp.1999.0734

Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence, 37,* 327–328. doi:10.1016/j.intell.2009.04.005

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18,* 46–60. doi:10.3758/s13423-010-0034-0

Nestor, P. G., & O'Donnell, B. F. (1998). The mind adrift: Attentional dysregulation in schizophrenia. In R. Parasuraman (Ed.), *The attentive brain* (pp. 527–546). Cambridge, England: MIT Press.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7,* 75–79. doi:10.1038/nn1165

Owens, M., Koster, E. H., & Derakshan, N. (2013). Improving attention control in dysphoria through cognitive training: Transfer effects on working memory capacity and filtering efficiency. *Psychophysiology, 50,* 297–307. doi:10.1111/psyp.12010

Passolunghi, M. C., Cornoldi, C., & De Liberto, S. (1999). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition, 27,* 779–790. doi:10.3758/BF03198531

Perkins, D., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18,* 16–25. doi:10.3102/0013189X018001016

Perry, J. L., & Carroll, M. E. (2008). The role of impulsive behavior in drug abuse. *Psychopharmacology, 200,* 1–26. doi:10.1007/s00213-008-1173-0

Plomin, R., & Spinath, F. M. (2002). Genetics and general cognitive ability (*g*). *Trends in Cognitive Sciences, 6,* 169–176. doi:10.1016/S1364-6613(00)01853-2

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Pædagogiske Institut.

Raven, J. C. (1990a). *Advanced Progressive Matrices Set I*. Oxford, England: Oxford Psychologists Press.

Raven, J. C. (1990b). *Advanced Progressive Matrices Set II*. Oxford, England: Oxford Psychologists Press.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., . . . Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142,* 359–379. doi:10.1037/a0029082

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition, 56,* 129–140. doi:10.1016/j.bandc.2004.09.016

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78,* 122–135. doi:10.1037/0022-3514.78.1.122

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566–594. doi:10.1037/0096-3445.132.4.566

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50,* 245–276.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138,* 628–654. doi:10.1037/a00274732012-06385-001

Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of go/no-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia, 46,* 224–232. doi:10.1016/j.neuropsychologia.2007.07.015

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 6791–6792. doi:10.1073/pnas.0803396105

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., . . . Kawashima, R. (2010). Training of working memory impacts structural connectivity. *Journal of Neuroscience, 30,* 3297–3303. doi:10.1523/jneurosci.4611-09.2010

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

Van der Molen, M. J., Van Luit, J. E. H., Van der Molen, M. W., Klugkist, I., & Jongmans, M. J. (2010). Effectiveness of a computerized working memory training in children with mild intellectual disabilities. *Journal of Intellectual Disability Research, 54,* 433–447.

Verbruggen, F., & Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience & Biobehavioral Reviews, 33,* 647–661. doi:10.1016/j.neubiorev.2008.08.014

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070. doi:10.1037/0022-3514.54.6.1063

Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Östensson, M. L., Bartfai, A., & Klingberg, T. (2007). Computerized working memory training after stroke—A pilot study. *Brain Injury, 21,* 21–29. doi:10.1080/02699050601148726

Westerberg, H., & Klingberg, T. (2007). Changes in cortical activity after training of working memory—A single-subject analysis. *Physiology & Behavior, 92,* 186–192. doi:10.1016/j.physbeh.2007.05.041

Zinke, K., Zeintl, M., Eschen, A., Herzog, C., & Kliegel, M. (2012). Potentials and limits of plasticity induced by working memory training in old-old age. *Gerontology, 58,* 79–87. doi:10.1159/000324240000324240

Appendix

Overall Effects of Between-Group Differences in Task Performance Comparing Pretest (Time 1) With Posttest (Time 2) and Follow-Up Testing (Time 3)

Variable	Time 1 versus Time 2			Time 1 versus Time 3		
	$F_{(df,\ \text{error})}$	p	η_p^2	$F_{(df,\ \text{error})}$	p	η_p^2
Go/no-go (training tasks)						
RT						
Group	$1.44_{(2,\ 116)}$	.240	.02	$0.98_{(2,\ 105)}$	.379	.02
Time × Group	$14.77_{(2,\ 116)}$	<.001	.20	$10.48_{(2,\ 105)}$	<.001	.17
d'						
Group	$1.12_{(2,\ 116)}$	.330	.02	$0.35_{(2,\ 105)}$	.703	.01
Time × Group	$0.20_{(2,\ 116)}$	.818	<.01	$0.25_{(2,\ 105)}$	.777	.01
cE%						
Group	$0.87_{(2,\ 116)}$	.421	.02	$0.13_{(2,\ 105)}$	.880	<.01
Time × Group	$1.55_{(2,\ 116)}$	.216	.03	$3.27_{(2,\ 105)}$	.042	.06
Stop-signal (training tasks)						
RT						
Group	$0.72_{(2,\ 114)}$	.488	.01	$0.63_{(2,\ 105)}$	.535	.01
Time × Group	$10.46_{(2,\ 114)}$	<.001	.16	$7.36_{(2,\ 105)}$	.001	.12
SSRT						
Group	$0.47_{(2,\ 114)}$	.629	.01	$2.43_{(2,\ 105)}$	.093	.04
Time × Group	$0.21_{(2,\ 114)}$	.812	<.01	$0.64_{(2,\ 105)}$	.529	.01
cE%						
Group	$0.11_{(2,\ 114)}$	.900	<.01	$0.01_{(2,\ 105)}$	.987	<.01
Time × Group	$3.51_{(2,\ 114)}$	.033	.06	$2.94_{(2,\ 105)}$	.057	.05
Stroop (near-transfer task)						
RT						
Group	$0.17_{(2,\ 114)}$	.842	<.01	$0.04_{(2,\ 107)}$	.961	<.01
Time × Group	$2.08_{(2,\ 114)}$	.130	.04	$1.22_{(2,\ 107)}$	.298	.02
Stroop effect						
Group	$1.13_{(2,\ 114)}$	.327	.02	$2.04_{(2,\ 107)}$	.135	.04
Time × Group	$1.68_{(2,\ 114)}$	.190	.03	$.51_{(2,\ 107)}$	.604	<.01
Fluid intelligence (far-transfer task)						
Group	$1.93_{(2,\ 117)}$	.150	.03	$2.50_{(2,\ 109)}$	.087	.04
Time × Group	$0.28_{(2,\ 117)}$	.756	<.01	$0.67_{(2,\ 109)}$	.513	.01

Note. RT = mean reaction time on correct trials; SSRT = Stop-Signal reaction time; cE% = commission error rate.

Received April 4, 2013
Revision received December 17, 2013
Accepted January 10, 2014 ■