

Attention Training in Schoolchildren Improves Attention but Fails to Enhance Fluid Intelligence

Justyna Sarzyńska, Dorota Żelechowska
The University of Social Sciences and Humanities

Marcel Falkiewicz
Nencki Institute of Experimental Biology,
Polish Academy of Sciences
Max Planck Institute for Cognitive and
Brain Sciences, Leipzig

Edward Nęcka
Jagiellonian University in Krakow
The University of Social Sciences
and Humanities

Fluid intelligence is a critical factor in learning and instruction. It also influences performance at school and in the workplace. There have been many attempts to directly and indirectly improve general fluid intelligence by training its underlying cognitive functions, such as working memory, cognitive control, or attention. The aim of the present study was to determine the extent to which school-age children's scores on intelligence tests could be improved by attention training. After training sessions, which consisted of four computerized cognitive tasks that practiced various aspects of attention, the children's scores on an attention test improved, with fewer false alarms and increased performance speed. This improvement partially persisted over an extended period of time. However, this effect was not associated with higher intelligence test scores. These results suggest that attention is possible to develop through short-term interventions but general intelligence is not. We interpret our findings in terms of the three-stratum theory of human intelligence.

Key words: attention, intelligence, cognitive training

Attention is at the core of cognition (Posner & Petersen, 1990; Petersen & Posner, 2012) and fluid intelligence is the most fundamental cognitive ability (Gottfredson, 1997; Jensen, 1998). Therefore, it is tempting to check whether both attention and intelligence can be improved through planned interventions. Being its cognitive substrate, some aspects of attention are

believed to underlie general mental ability (Schweizer, Moosbrugger, & Goldhammer, 2005; Stankov, 1988). If so, the effects of attention training should generalize to fluid intelligence. Here we attempt to verify this hypothesis with the participation of young schoolchildren.

Fluid intelligence is understood to involve general reasoning ability, problem solving skills, and abstract thinking in novel situations (Cattell, 1971). The distinction between fluid and crystallized intelligence has been proposed by Raymond Cattell (1957, 1971) and later developed by John Horn (1968). According to Cattell, fluid intelligence (Gf) is a biologically determined "pure" ability to reason in the inductive or deductive way, whereas crystallized intelligence (Gc) is a culturally determined ability to acquire and use knowledge. John Horn (1968;

Preparation of this paper was supported by the grant No. N N106 328239 from the Polish Ministry of Science and Higher Education.

Correspondence concerning this paper should be addressed to: Justyna Sarzyńska, University of Social Sciences and Humanities, ul. Chodakowska 19/31, 03-815 Warszawa, Poland.
E-mail: jsarzyńska@swps.edu.pl

Received December 24, 2015

Horn & Stankov, 1982) developed the Cattell's distinction, adding several lower-level abilities, such as visual processing (Gv), short-term memory skills (Gsm), long-term memory skills (Glr), or speed of processing (Gs). As we can see, all acronyms contain the letter G, meaning that these abilities are relatively general in nature, that is, they manifest themselves in numerous and diverse tasks and situations rather than in some specific settings. John Carroll (1993, 1997) developed the Cattell-Horn theory into the form of a three-stratum theory of intelligence. According to his approach, human abilities are organized in three strata, depending on the number and variety of tasks and situations in which they can be observed. The most general ability is Spearman's (1927) general factor *g*, representing the highest stratum III. The second level (stratum II) includes so-called "broad abilities", whose number was originally eight but today it ranges from eight to sixteen, depending on the version of the model. The most important broad abilities are *Gf* and *Gc*, already proposed by Cattell and Horn. Other II stratum abilities are, for instance, processing speed or retrieval ability. The lowest level (stratum I) includes abilities that manifest themselves in very specific, narrowly defined tasks. They may also depend on specific cognitive abilities adopted by a person. Carroll's conceptualization of Cattell and Horn's ideas is now recognized as the CHC (Cattell-Horn-Carroll) theory of intelligence. It is regarded to be the most comprehensive model of the structure of human abilities, supported by confirmatory factor-analytical studies (e.g., Flanagan & Dixon, 2014; Gustafsson & Undheim, 1996).

Ideally, it would be suitable to investigate training effects in reference to all the abilities described by the CHC theory but such an agenda is hard to implement in a single study. Besides, we do not have appropriate, theory-based tools for the measurement of potential training effects. For these reasons, we decided

to focus on an ability that is general enough and predicts significant life achievements, namely, the general fluid intelligence (*Gf*). It has been demonstrated that *Gf* is an important predictor of academic achievement (Deary, Strand, Smith, & Fernandes, 2007), career outcomes and professional achievement (Ree & Earles, 1992), as well as health and mortality (Deary, 2008; Gottfredson, 1997, 2004). Not surprisingly, many planned interventions that include cognitive training have sought to improve it. However, it is not clear whether fluid intelligence is susceptible to such interventions. Many studies have found that intelligence training yields only limited effects (Barnett & Ceci, 2002). According to Sternberg (2008), the training tasks used in many studies that claim to improve fluid intelligence were very similar to the final tests that measured training effects (e.g., Kramer & Willis, 2002), which suggests that the training improved only specific test skills. Although intelligence test scores can be improved by practicing test taking, this effect does not contribute to our understanding of the nature of intelligence and has limited practical value.

Another line of research shows that improvement in fluid intelligence may result from transfer effects when certain basic cognitive functions are practiced, such as working memory (Jaeggi, Buschkuhl, Jonides, & Perrig, 2008; Jaeggi, Buschkuhl, Jonides, & Shah, 2011; Klingberg et al., 2005), executive functions (Schweizer, Hampshire, & Dalgleish, 2011), or attention (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005). Arguments supporting the hypothesis that this type of training can boost intelligence are as follows. First of all, many studies indicate that elementary cognitive functions underlie more complex ones, such as intelligence (Hunt, 1980; Sternberg, 1985). Secondly, elementary and complex functions tend to have similar brain substrates (e.g., Duncan, 2003; Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010;

Shimamura, 2000). Therefore, enhancement of simpler functions should indirectly affect more complex ones.

Some functions of attention are believed to be the cognitive substrate of intelligence (Nęcka, 1996; Stankov, 1988). For example, Coull's model (1998) distinguishes four aspects of attention: orienting attention, divided attention, selective attention, and vigilance. All of them seem to be related to fluid intelligence (Crawford, 1991; Nęcka, 1996; Roberts, Beh, & Stankov, 1988; Roberts, Beh, Spilsbury, & Stankov, 1991; Rockstroh & Schweizer, 2004; Schweizer, 2001; Stankov, 1988). There is also a common neural substrate responsible for the association between attention and intelligence. Duncan and Owen (2000) found that tasks measuring attention (e.g., Stroop) and intelligence (e.g., problem solving) both involve activity in the following brain areas: the dorsolateral and ventrolateral prefrontal cortices, and the dorsal anterior cingulate cortex (see also: Shimamura, 2000).

Rueda et al.'s (2005) study yielded support for the hypothesis that attention training can enhance intelligence. Four- and six-year-old children performed nine (four-year-olds) or ten (six-year-olds) computerized tasks training different aspects of executive attention (Posner & Petersen, 1990) over a period of two to three weeks. Effectiveness was evaluated with the children's version of the Attention Network Test (Rueda et al., 2004) and the Kaufman Brief Intelligence Test (Kaufman & Kaufman, 1990). After training, children's performance improved on both behavioral and neural indicators. EEG data showed that training yielded effects similar to those associated with brain maturation, particularly in the six-year-olds. Their brain activity patterns became similar to those observed in adults. Intelligence test results also improved, particularly in the components that assessed fluid intelligence. Karbach and Kray (2009) conducted another study that demonstrated far

transfer effects on intelligence. The authors applied four different versions of task-switching training to healthy participants in three different age groups (8-10, 18-26, and 62-76 years). They checked for near transfer using different versions of tasks requiring switching, and far transfer to working memory, inhibitory control, and fluid intelligence. Regardless of the age group, participants who trained task-switching improved their performance in other tasks, including intelligence tests. Thus, Karbach and Kray (2009) demonstrated both near transfer effects, with criterial tasks similar to training tasks, and far transfer effects, with criterial tasks not resembling the training tasks (see: Shipstead, Redrick, & Engle, 2010).

Tucha et al. (2011) also demonstrated effective training of attention. Thirty-six children diagnosed with ADHD and 16 healthy children (passive control group) participated in the study. The children with ADHD were divided into two experimental groups that trained either attention or perception. Trainings were held twice a week for two months. Criterial tests consisted of six computer tasks measuring various aspects of attention (Zimmermann & Fimm, 2002). Children who underwent attention training exercised vigilance, selectivity, and divided attention, according to the AixTent program (Sturm, Orgass, & Hartje, 2001). Children from the second experimental group performed the German version of the Frostig Developmental Test of Visual Perception (Frostig, Horne, & Miller, 1972; Reinartz & Reinartz, 1974). The results showed that children from the first experimental group improved their scores in tasks that measured vigilance, flexibility, and divided attention. Such changes were not found in healthy control children or children who practiced perceptual tasks, indicating that the improvement did not stem from mere repetition of assessment with the same tasks.

Tucha et al.'s (2011) study demonstrated both near and far transfer on previously untrained

aspects of attention, but the author did not measure transfer to general intelligence. The current research therefore aimed at application of training tasks similar to those used in Tucha et al.'s (2011) study but also referring to all four aspects of attention highlighted by Coull (1998). In contrast to Tucha et al.'s study, we tested healthy children and controlled for possible delayed effects of training. We also attempted to determine whether fluid intelligence could be improved through attention training with limited engagement of executive functions. Attention, executive control, and working memory are closely related concepts (Friedman et al., 2006; Klauer & Phye, 2008; Shimamura, 2000; Unsworth & Engle, 2006). Hence, training effects are likely to be contaminated and misinterpreted. In this study, we deliberately aimed at training attention with limited engagement of executive control.

Method

Participants

Fifty-four eight-year-old elementary school children participated in the study, with 30 (14 male) in the experimental group and 24 (9 male) in the control group. One child with clinical diagnosis of developmental disorders was excluded from the study. Twenty-seven children in the experimental group and 23 in the control group met the compliance criteria of participation in 10 training sessions and all criterial tests. All participants took tests measuring intelligence and attention, and then participated in either attention training (experimental group) or another cognitive training (control group). Tests were administered before (*pretest*) and after all training sessions (*posttest*). Additionally, the experimental group was assessed for the third time three months after the *posttest* (*follow-up test*), in order to establish the stability of training effects. Children from both groups com-

pleted 10 training sessions that were held two to three times a week and lasted approximately 30 minutes. During each session, children performed all training tasks, which means that they spent approximately 7-8 minutes on each task in the experimental group (4 tasks) and 15 minutes in the control group (2 tasks). The tasks always appeared in random order.

Criterion Measures

The Standard Edition Raven's Progressive Matrices test (RPM; Raven, Raven, & Court, 2003) was used to assess general intelligence because it is regarded to provide the best approximation of the general reasoning ability (Snow, Kyllonen, & Marshalek, 1984). RPM consists of 60 questions of increasing difficulty. Each question presents a matrix of patterns in which one pattern is missing. The task is to select the missing pattern among a set of given alternatives. We used the classical form of RPM for the pretesting and the parallel version after the training sessions.

Attention was measured using the d2 Attention test (Brickenkamp & Zillmer, 1998), which assesses the overall efficiency of attention. Participants were given a sheet of A4 paper containing fourteen lines of letters and were instructed to cross out properly labeled "d" letters in rows of variously marked "d" and "p" letters. The task was timed, and participants were allowed 20 seconds to complete each row so that they had just under five minutes to complete the entire test. This test provides measures of perceptual speed, the overall perceptual ability (the rate ratio adjusted to the number of errors), hits (which reflects concentration) and the number and types of errors (misses or false alarms). In addition, two measures based on the Signal Detection Theory (SDT, Green and Swets, 1966) were added: d' reflects the ability to distinguish signal from noise, whereas c' provides the bias measure. Higher values of

c' indicate a bias towards the conservative strategy (more correct rejections with greater number of omissions), and lower values indicate the liberal strategy (increased number of hits at the expense of higher frequency of false alarms).

Experimental Training

Four computer games that trained different aspects of attention were designed for the study. The "Fish" task required participants to distinguish signals from noise. There were several types of fish swimming in the aquarium, and children had to operate with a carnivorous fish that was "eating" fish of one type while ignoring other species. Once 30 fishes were eaten, the rules changed: children were required to inhibit their formerly acquired habit and "eat" a different type of fish. The difficulty of the task was adjusted to the child's current level of performance. Training progress was facilitated or hindered by changing velocity of the fishes, the number of edible fishes, and velocity of the fish controlled by the child. Additionally, the initial game could exhibit one of three levels of difficulty, differing in the ease with which edible fishes could be discerned (e.g., shape and color, shape alone, or shape visible only in the stream of light switched on by the child). Total number of wrong fishes "eaten" during the session was used as a performance measure.

The "Easter eggs" task, which trained alertness, was modeled on the classic Mackworth task (1948). In this task, children controlled the work of a bunny in the Easter egg factory. Colored Easter eggs moved along a conveyor belt, and the bunny's task was to reject damaged eggs or eggs that were not completely colored. The damage on the eggs became less visible over time, although the rate at which they moved along the conveyor belt remained constant. Total number of incorrect responses was used as a performance indicator.

The "Apples" task, which assessed orienting attention, was based on Posner's cuing paradigm (1978). Children were required to focus on a fixation point flanked by two boxes and control the action of a wolf collecting apples that appeared on both sides of the fixation point. The task was to press a key as quickly as possible when an apple appeared inside of one of the boxes. Additionally, the appearance of apples was preceded by a cue (appearance of twigs to one side), although the cues were sometimes misleading. Sometimes the child had to respond to an apple appearing in the location where a cue had previously appeared (saccades), and sometimes to an apple appearing on the opposite side (antisaccades). During the game, the proportion of correct cues, and the interval between the appearance of the cue and the stimulus, decreased. The dependent variable for this task was the total number of incorrect responses for each session.

The "Jigsaw" task, which trained divided attention, involved simultaneously playing two puzzles. An item would appear at the center of the screen, which the child had to assign to one of two images that were presented simultaneously. We manipulated the difficulty of the task by dividing the images into an increasing number of elements, increasing similarity of the presented images, and shortening the period of time within which the child could respond. The total number of correct responses per session was used as a performance measure.

Control Training

Children in the control group participated in training with the same intensity and format as children in the experimental group, but the training tasks placed minimal demands on attention. Their training consisted of two tasks practicing problem solving or perceptual skills. In the "Mice Game", players went through a maze collecting pieces of cheese while avoiding ob-

stacles (such as mousetraps or hamster wheels) that would send them back to the beginning. The time to go through the maze was limited, and difficulty was manipulated by the complexity of the maze, the number of obstacles, and the amount of cheese to collect. The main goal of “Machine Space” was to build a machine from a number of given elements by dragging them into the right places. To eliminate the role of memory in the task, players saw a sketch of this machine. Time to complete the task was limited and difficulty was manipulated by the number of elements and their similarity. The tasks performed by children in the control group were as attractive as those used in the experimental group, as assessed in post-training interviews.

The same token system was introduced to enhance motivation in both the experimental and control groups. After each session, children could choose a sticker to be placed on a card that they had previously received. After earning five stickers (i.e., after completing half of the trainings) they received a prize – either a pen or a keychain. After earning 10 stickers (i.e., at the end of training), the children could choose a prize worth about 12 €.

Results

Mixed-effects ANOVA, with covariate measures of training performance, was used for sta-

tistical modeling. All analyses were carried out using R (R Core Team, 2015) with the STAN (Hoffman & Gelman, 2011) and several additional packages (supplementary material: [1] data: <http://goo.gl/T63DA8>, [2] statistical modeling: <http://goo.gl/FleUoy>).

Firstly, we present the effects of practice, that is, improvement of training tasks performance. Next, we examine near transfer effects for attention, using the following measures of the d2 test: perceptual speed, overall perceptual ability, hits, omissions, false alarms, discriminability (d'), and bias (c'). Finally, we present far transfer effect for fluid intelligence.

Practice Effects

Table 1 shows the improvement in training task performance (decreased number of errors and increased number of correct responses). We can see a comparison between the first and last training sessions. The relevant statistics are: $t(25) = 3.08, p = 0.002, d = 0.6$, for Task 1 (*Fish*), $t(25) = 2.74, p = 0.005, d = 0.54$, for Task 2 (*Easter eggs*), $t(26) = 3.15, p = 0.002, d = 0.61$, for Task 3 (*Apple*), and $t(26) = 4.66, p < 0.001, d = 0.89$, for Task 4 (*Jigsaw*). The effects presented in Table 1 justify the conclusion that the training procedures proved their effectiveness. It is therefore reasonable to expect possible near transfer and far transfer effects. Changes of performance across training sessions are presented in Figure 1.

Table 1 *Practice effects: The results (means and standard deviations) obtained by the experimental group in four training tasks, depending on the training session*

	1st session	10th session	<i>t</i> value 10th vs 1st session
Task 1. Fish ¹	0.25 (0.09)	0.21 (0.08)	-3.08**
Task 2. Easter eggs ¹	0.03 (0.05)	0.01 (0.01)	-2.74**
Task 3. Apples ¹	0.24 (0.21)	0.10 (0.1)	-3.15**
Task 4. Jigsaw ²	1.67 (21.43)	24.85 (20.48)	4.66***

Note. ¹ Ratio of the number of errors to the total number of responses.

² The number of correct responses minus the number of errors.

** $p < 0.01$, *** $p < 0.001$

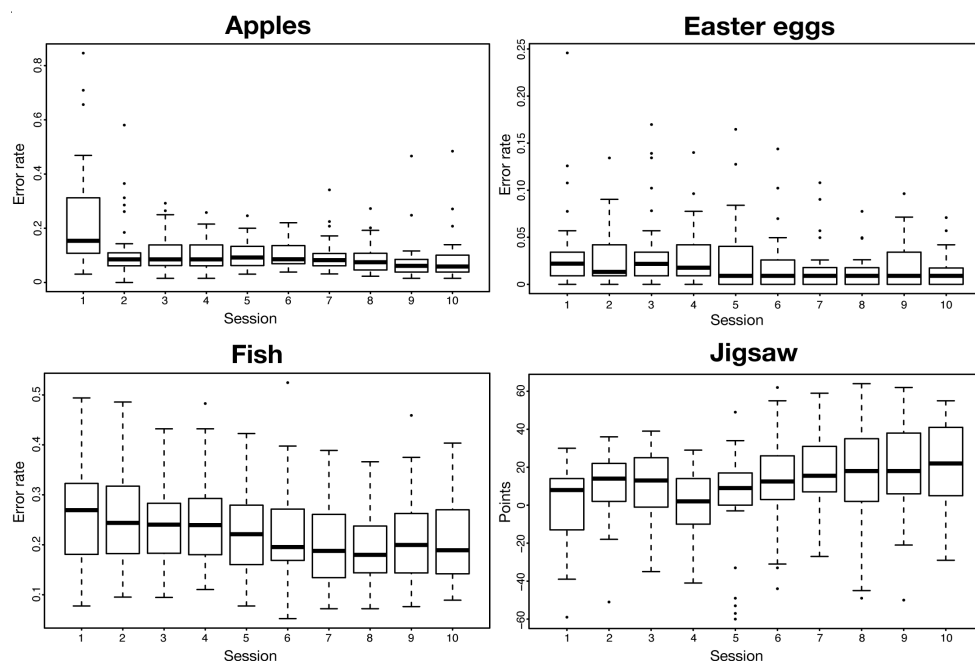


Figure 1 Changes of performance across training sessions

Solid horizontal bars in boxes denote the median. Top and bottom edges of the box represent 75% and 25% quartile, respectively. The top and bottom whiskers denote maximum and minimum values after excluding the outliers. The points above and below the whiskers represent outliers.

Near Transfer Effects: Attention

Results obtained by both groups in the d2 test and RPM are presented in Table 2.

A single-step multiple comparisons correction based on joint t distribution of the linear function (Hothorn, Bretz, & Westfall, 2008; Bretz, Hothorn, & Westfall, 2010) was applied for each dependent variable. Additionally, we provide confidence intervals for each comparison.

Perceptual Speed

The analysis revealed a significant main effect of measurement for the number of processed

items ($F(1,47) = 25.28, p < 0.0001, \eta^2 = .08$), indicating that both groups performed better in the posttest. There was no significant between-group difference for the pretest (0.34, 95% CI: [-43.24 – 43.93], $t = 0.02, p = 0.98$). After training sessions, the control group improved by 45.5 (95% CI: [19.48 – 71.52], $t = 4.4, p \leq .001$) and the experimental group by 24.63 (95% CI: [1.14 – 48.12], $t = 2.64, p = .02$). This improvement was not maintained for the experimental group in the follow-up test, as the number of processed items decreased by 39.89 (95% CI: [18.47 – 61.31], $t = 4.23, p = .0002$) compared to the posttest and decreased by 15.26 compared to the pretest (95% CI: [-36.68 – 6.16], $t = 1.62,$

$p = .11$). Lack of convincing between group differences suggests that improvement in processing speed was not specifically related to attention training.

Overall Perceptual Ability

A similar pattern was found for overall perceptual ability (Table 2). There was a significant main effect for measurement ($F(1,47) = 29.48, p < .0001, \eta^2 = 0.1$). Groups did not differ in the pretest scores (1.04, 95% CI: [-41.25 – 43.34], $t = 0.06, p = 0.95$). The control group improved by 44.09 (95% CI: [17.92 – 70.27], $t = 4.23, p < 0.001$) and the experimental group by 32.11 (95% CI: [8.49 – 55.74], $t = 3.42, p = .003$).

Follow-up tests did not show evidence for persistence of training effects in the experimental group, as the scores decreased by 35.67 compared to the posttest (95% CI: [15.29 – 56.04], $t = 3.98, p < .001$) and did not differ from pretests (-3.55, 95% CI: [-23.93 – 16.82], $t = 0.4, p = 0.69$). This effect also cannot be explained specifically by attention training.

Hits

There was a main effect of measurement for the number of hits (Table 2, $F(1,47) = 25.32, p < .0001, \eta^2 = 0.1$). The experimental group had 15.21 more hits in the pretest (95% CI: [-4.7 – 35.12], $t = 1.92, p = 0.11$) compared to the control group.

Table 2 Descriptive (means and standard deviations) statistics concerning the performance measures

Performance measure	Group	Pretest	Posttest	Follow-up	<i>t</i> value pre- vs. posttest	<i>t</i> value pretest vs. follow-up
Perceptual speed	CTRL	237.55 (66.16)	283.05 (63.32)		4.4***	
	EXP	237.89 (58.79)	262.52 (54.31)	222.63 (30.91)	2.64*	-1.62
Overall perceptual ability	CTRL	220.36 (64.5)	264.45 (63.92)		4.23***	
	EXP	221.41 (52.4)	253.52 (54.89)	217.85 (30.87)	3.42**	-0.4
Hits	CTRL	75.27 (32.95)	97.5 (33.87)		4.12***	
	EXP	90.48 (20.08)	104.85 (23.63)	97.37 (12.31)	2.95**	2.01*
Omissions	CTRL	13 (11.74)	12.36 (10.97)		-0.26	
	EXP	11.78 (15.27)	7.33 (11.73)	0.89 (1.67)	-2.01	-4.31**
False alarms	CTRL	4.18 (5.4)	6.23 (7.71)		2.11	
	EXP	4.7 (4.71)	1.67 (2.59)	3.89 (2.61)	-3.47**	-1.02
Discriminability (d')	CTRL	3.22 (0.85)	3.35 (0.86)		0.79	
	EXP	3.46 (0.76)	4.35 (0.8)	4.45 (0.41)	6.00***	5.9***
Bias (c')	CTRL	0.48 (0.29)	0.37 (0.33)		-1.61	
	EXP	0.25 (0.36)	0.36 (0.33)	-0.3 (1.11)	1.85	-8.47***
RPM	CTRL	27.68 (8.07)	28.09 (9.13)		0.43	
	EXP	29.63 (7.21)	31.19 (7.83)	30.93 (6.63)	1.8	1.57

Note. Perceptual speed: The number of items scanned during the five minutes period. Overall perceptual ability: The number of scanned items minus the number of errors (interpreted as an index of overall efficiency of attention). Hits: The number of correctly spotted signals (properly labeled 'd' letters). Misses: The number of missed signals. False alarms: The number of erroneously detected non-signals (other than properly labeled 'd' letters). Discriminability (d'): ability to discriminate between signal and noise. Bias: response strategy – lower values indicate more conservative approach. RPM: Raven's Progressive Matrices score.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Number of hits increased by 22.23 (95% CI: [8.69 – 35.76], $t = 4.12$, $p < 0.001$) in the control group and by 14.37 (95% CI: [2.15 – 26.59], $t = 2.95$, $p = 0.009$) in the experimental group. Follow-up tests revealed that the number of hits remained 6.89 higher (95% CI: [-1.22 – 15], $t = 1.93$, $p = 0.07$) compared to the pretest and dropped by 7.48 (95% CI: [-0.63 – 15.59], $t = 2.01$, $p = 0.07$) compared to the posttest. Though the difference in pretest scores necessitates caution with interpretation, analyses indicate a specific effect of attention training on the number of detected items in the d2 test. Moreover, this effect persisted over an extended period of time to the follow-up test.

Omissions

Mixed-effects ANOVA revealed no significant main effects and interactions for the number of omissions (Table 2). However, a significant effect of measurement ($F(1,68, 43.55) = 9.37$, $p = .0008$, $\zeta^2 = 0.14$) was found for the experimental group only. The number of omissions in the experimental group was lower in the follow-up test by 10.89 (95% CI: [5.14 – 16.64], $t = 4.31$, $p = .0001$) compared to the pretest and by 6.44 (95% CI: [0.69 – 12.19], $t = 2.55$, $p = 0.01$) compared to the posttest. As we can see, the number of omissions in the experimental group decreased after training, which may be an effect specific for attention training.

False Alarms

There was a significant group \times measurement interaction (Table 2, $F(1,47) = 15.14$, $p = .0003$, $\eta^2 = 0.06$) for the number of false alarms, without significant main effects. The difference between groups in the pre-test was negligible (.52, 95% CI: [-3.28 – 4.52], $t = 0.35$, $p = 0.73$). The experimental group decreased the number of false alarms in the posttest by 3.04 (95% CI: [0.84 – 5.24], $t = 3.47$, $p = 0.002$), but the control

group increased the number of false alarms by 2.05 (95% CI: [-0.39 – 4.48], $t = 2.11$, $p = 0.08$). Contrasts between groups showed a larger effect in the experimental group by 5.08 (95% CI: [1.8 – 8.36], $t = 3.89$, $p < 0.001$). Follow-up tests showed an increase in false alarm rate compared to the posttest by 2.22 (95% CI: [0.4 – 4.04], $t = 2.78$, $p = 0.015$) and was only slightly lower than for the pretest (0.81, 95% CI: [-1 – 2.63], $t = 1.01$, $p = 0.31$). This result indicates that attention training had a specific positive effect on the number of false alarms.

Discriminability and Bias

The decrease in the frequency of false alarms and omissions, coupled with the increase in the information processing speed and hits, suggests that children who underwent attention training were more careful during the task. The control group participants also improved information processing speed and the number of hits, but at the expense of more false alarms. We interpret these findings in terms of SDT's (*Signal Detection Theory*, Green & Swets, 1966) constructs of discriminability (d') and bias (c'). Both parameters were calculated using a Bayesian hierarchical model (Lee & Wagenmakers, 2013), separately for each group and measurement. Maximum a posteriori estimates were further analyzed using mixed-effects ANOVA.

The results are presented in Table 2. The analysis revealed a significant main effect of group for discriminability ($F(1,47) = 9.13$, $p = .004$, $\eta^2 = 0.13$), a main effect of measurement ($F(1,47) = 21.23$, $p < 0.0001$, $\eta^2 = 0.09$), and a group \times measurement interaction ($F(1,47) = 11.79$, $p = 0.001$, $\eta^2 = .05$). The difference between experimental and control groups for the pretest was small: d' was 0.24 larger in the experimental group (95% CI: [-0.34 – 0.83], $t = 1.04$, $p = .5$). The index d' did not change after training in the control group (0.13, 95% CI: [-0.28 – 0.54], $t = 0.79$, $p = 0.5$), but increased in the ex-

perimental group by 0.89 (95% CI: [0.52 – 1.26], $t = 5.99, p < .0001$). The increase in experimental group was 0.76 higher (95% CI: [0.21 – 1.31], $t = 3.43, p = .002$) than in the control group. The effect of training was maintained in the follow-up test. The index d' remained .99 higher (95% CI: [0.61 – 1.38], $t = 5.89, p < .0001$) compared to pretest and was higher by .11 (95% CI: [-0.28 – 0.49], $t = 0.63, p = 0.53$) compared to the posttest. Means and confidence intervals for d' are provided in Figure 2. These results indicate that there was a specific effect of attention training on the ability to distinguish signals from noise in the d2 test (see: Figure 2).

For bias, there was a significant group \times measurement interaction (Table 2, Figure 3, $F(1,47) = 5.91, p = 0.02, \eta^2 = 0.03$). Bias was higher in the pretest for the control group compared to the experimental group by 0.23 (95% CI: [0 – 0.46], $t = 2.41, p = .06$). After training, bias decreased in the control group by 0.11 (95% CI: [-0.06 – 0.29], $t = 1.61, p = 0.11$) but increased in the experimental group by 0.12 (95% CI: [-.04 – .27], $t = 1.85, p = 0.07$). The difference between

experimental and control groups was 0.23 (95% CI: [0 – .46], $t = 2.43, p = 0.06$).

Follow-up tests in the experimental group showed a large decrease in bias for the experimental group compared to both the pretest (0.55, 95% CI: [0.4 – 0.69], $t = 8.48, p < .0001$) and the posttest (0.67, 95% CI: [0.52 – 0.81], $t = 10.29, p < 0.001$). Means and standard deviations for bias are shown in Figure 3. In addition to better discrimination of signal and noise, attention training shifted the strategy towards a more conservative approach. The control group showed the opposite effect in terms of strategy: it became more liberal. However, bias in the follow-up indicates that participants became even more liberal than in the pretest (see: Figure 3).

Far Transfer Effects: Fluid Intelligence

Table 2 also demonstrates the results obtained by two groups in the intelligence test and Figure 4 illustrates these findings. No significant training effects were revealed for RPM scores,

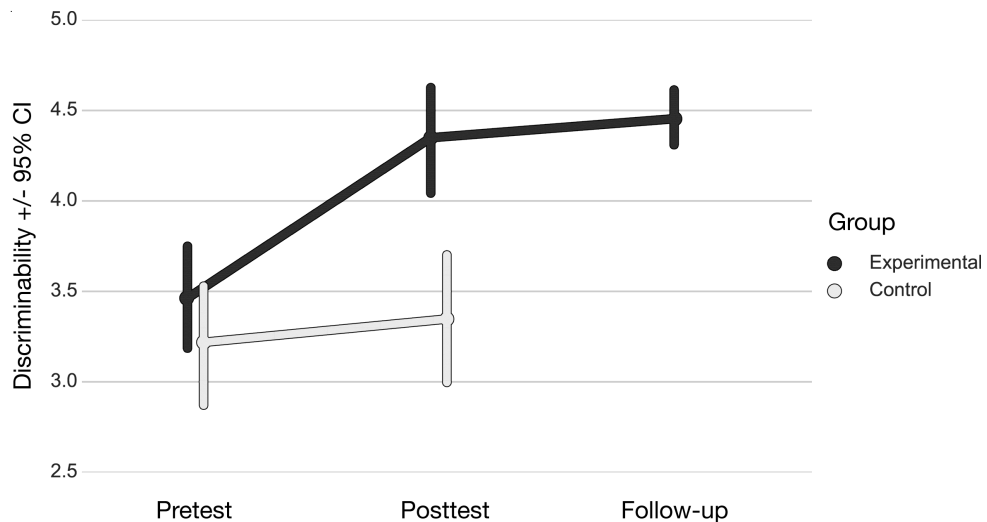


Figure 2 The value of d' depending on the group and the time of measurement

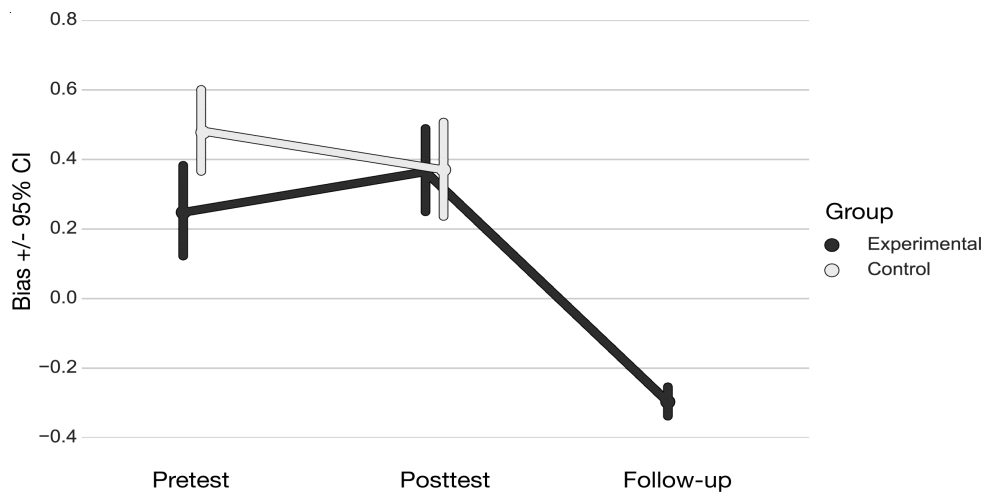


Figure 3 The value of bias depending on the group and the time of measurement

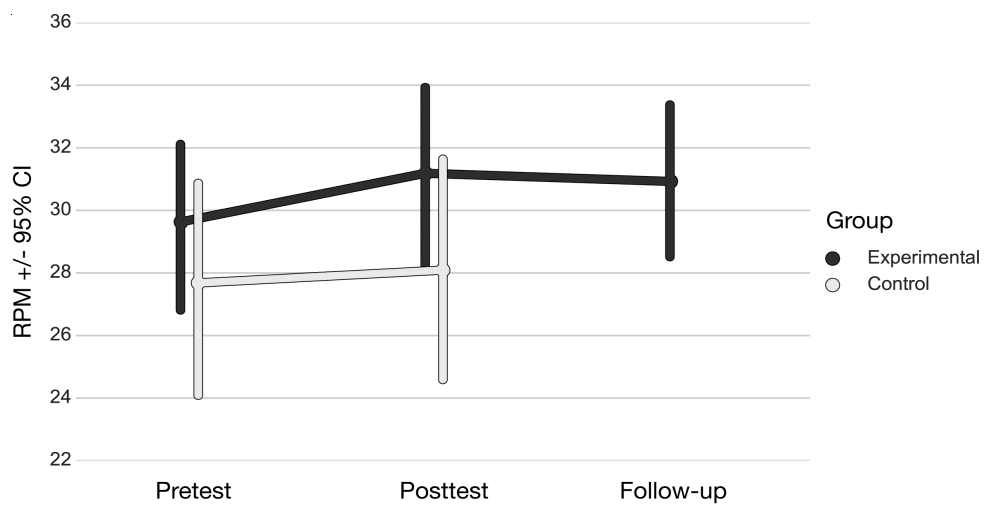


Figure 4 Raw scores in Raven's Progressive Matrices depending on the group and the time of measurement

with negligible improvement in both groups (main effects: $F(1,47) = 1.3, p = 0.26, \eta^2 = 0.02$, and $F(1,47) = 2.33, p = 0.13, \eta^2 = 0.004$, for group and measurement, respectively; interaction: $F(1,47) = 0.79, p = 0.38, \eta^2 = 0.001$). However, there seems to be a slight (although non-significant) increase in the posttest RPM score for the experimental group. This improvement seems to persist in the follow-up test. In order to make sure that the training intervention did not result in enhancement of intelligence, we performed a Bayesian analysis and computed Bayes Factors (BF), which is a method to estimate the likelihood of the null hypothesis (Kass & Raftery, 1995). BF greater than one indicates that the null hypothesis gained greater probability than the alternative one. Concerning the pretest-posttest difference in the experimental group, BF obtained the value 1.0, indicating that the null hypothesis (i.e., training does not work) should not be rejected, although it appeared exactly as probable as the alternative hypothesis (i.e., training works). However, BF concerning the pretest vs. follow-up test obtained the value 1.2, which means that the null hypothesis appeared slightly more likely than the alternative one. As to the control group, the BF obtained the value 5.0, which means that the null hypothesis was five times more likely than the alternative hypothesis. Altogether, we interpret these findings in terms of inefficiency of attention training for enhancement of general fluid intelligence (Gf).

Discussion

We found that children in the experimental group improved their performance in the attention test, increasing the number of analyzed items and detected signals, while decreasing the number of false alarms and omissions. Moreover, the SDT analysis confirmed that children from the experimental group changed their strategy, by becoming more careful in the d2 task. They

also exhibited higher discriminability scores, indicating improved ability to differentiate signals from noise. Three months following the completion of training, their processing speed went down and the false alarm rate returned to initial levels, but the higher number of hits and lower number of omissions persisted. As for the SDT indices, improved discriminability was maintained but the strategy returned to a less careful one. So, the benefits of attention training persisted in only one aspect of the d2 test performance.

The improvement of attention observed immediately after training did not extend to children's intelligence test scores. This absence of far transfer probably cannot be explained by duration of training, because in other studies (e.g., Rueda et al., 2005) effects were observed after an even shorter intervention. The children's age does not seem to be crucial, either. Although there is much evidence suggesting that training may be more effective in the case of younger children (Wass, Scerif, & Johnson, 2014), there are studies demonstrating enhancement of intelligence in young and older adults as well (e.g., Karbach & Kray, 2009).

One possible explanation is that the relationship between intelligence and attention is not as direct, or as strong, as in the case of other functions. This conclusion is consistent with previous studies reporting that working memory is strongly related to intelligence, while other cognitive functions, such as attention, are associated with measures of intelligence only modestly or not at all (e.g., Engle, Tuholski, Laughlin, & Conway, 1999). Moreover, in most studies demonstrating the existence of a relationship between attention and intelligence, executive functions, such as inhibition, attention shifting, and working memory updating, were not assessed separately for their effect on intelligence test scores. Friedman et al. (2006) addressed this issue in a study of 234 individu-

als who performed nine tasks, three for each cognitive function, that are documented in the literature to assess inhibitory control, attention shifting, and working memory updating. Factor analysis indicated that for both fluid and crystallized intelligence only the association with memory updating was significant, explaining 41% to 48% of the variance. Other cognitive functions explained only 2% to 14% of the variance. Mogle, Lovett, Stawski, and Sliwinski (2008) also hypothesized that attention is not related to intelligence, and that the only important factor is the ability to accurately extract information from memory. In addition, Unsworth and Engle (2006) believed that the results that were previously interpreted as supporting the attention-intelligence relationship might instead result from misinterpretation of working memory indices. These hypotheses are supported by the data from trainings, in which improvements in fluid intelligence usually took place after exercising executive functions and working memory. It is worth noting that both Karbach and Kray's (2009) and Rueda et al.'s (2005) trainings involved executive functions, which may be more important for intelligence than the attentional abilities investigated in our experiment.

Our results seem compatible with the three-stratum theory of intelligence (Carroll, 1993, 1997; see also: Jensen, 1998). We definitely did not find any evidence that the g-factor (stratum III) might be enhanced through planned training of attention. Broad abilities (stratum II) were not affected by training, either, since this layer includes general fluid intelligence (Gf), assessed with Raven's matrices; we did not observe any transfer effects. Thus, we conclude that our training games enhanced narrow abilities (stratum I), in particular, the ability to concentrate attention and to avoid careless production of false alarms. Some authors (e.g., te Nijenhuis, Vianen, & van der Flier, 2007) claim that cognitive training does not in fact affect the third

stratum but only the lowest level of the hierarchy, consisting of specific abilities that closely resemble the trained skills. In our study, there was not much similarity between training games and criterial tests, so we conclude that our participants improved their narrow cognitive abilities rather than just specific testing skills. However, the problem of task specificity in cognitive training must be taken seriously (see: Klauer & Phye, 2008; Stankov, 1986; Herrnstein, Nickerson, Desanchez, & Swets, 1986).

Despite the failure to observe any improvement in fluid intelligence, enhancement of attention after only five hours of practice is a promising result. Since attention deficits are common in children and primarily manifest as neurodevelopmental disorders such as ADHD (Bulut, 2005), it is important to develop effective tools to support and enhance this cognitive function. Besides, attention is particularly important for children's academic achievement. Duncan et al. (2007), who analyzed the results of six longitudinal studies of children and adolescents, found that the strongest predictors of academic outcomes were: 1) the ability to focus attention and 2) basic academic skills, such as counting and reading. Our findings suggest that, in general, training children's attention is possible and worthwhile.

It is worth noting that the analysis of discriminability and bias parameters allowed us to pinpoint that children discriminated signals from noise more easily after attention training, while simultaneously promoting a more cautious strategy. The observed change in strategy is noteworthy because cognitive strategies play an important role in human intellectual activity, including intelligence (Hunt, 1980). In the present study, the change in strategy did not affect fluid intelligence. However, adoption of a strategy that involves a decreased tendency to commit false alarms may help to exert inhibitory control of impulsive reactions. This might prove useful in school, where it is often necessary to fully

consider the available options before responding. Adopting a “careful” strategy can therefore improve children’s use of the knowledge that they already possess.

As to limitations of this study, the five-hour training might be considered as relatively short. Next, a larger and heterogeneous transfer task battery would be advisable, in order to identify the specific aspects of attention that change due to training. Also, the concept of intelligence is broad and heterogeneous, including the aspects that were totally neglected in our study, such as practical intelligence or solving ill-defined problems (Ruisel, 2003). Lack of randomization, forced by practical reasons, is also a significant limitation of this study, but this is the case of many studies of cognitive trainings. The change in strategic behavior would need to be evaluated concerning its generalization across different tasks. Furthermore, the “careful” type of strategy is typically associated with academic outcomes, so in future studies it might be advisable to use a test of crystallized intelligence and to control for academic performance before and after training. The application of selected subscales of the Wechsler Intelligence test might also provide more conclusive results (McDougall & House, 2012). Moreover, it would be worthwhile to introduce tasks that could measure the extent to which the practiced skills transfer to everyday situations.

Altogether, we believe that attention is susceptible to improvement through planned interventions, although far transfer effects to fluid intelligence seem hard to corroborate.

References

- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*(4), 612-637.
- Bretz, F., Hothorn, T., & Westfall, P. (2010). *Multiple comparisons using R*. CRC Press.
- Brickenkamp, R., & Zillmer, E. (1998). *The d2 test of attention*. Hogrefe & Huber Pub.
- Bulut, S. (2005). Classroom interventions for children with attention deficit hyperactivity disorder. *Studia Psychologica*, *47*, 51-59.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Carroll, J. B. (1997). The three-stratum theory of cognitive abilities. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 122-130). New York, NY: Guilford Press.
- Cattell, R. B. (1957). *Personality and motivation structure and measurement*. New York, NY: World Book.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. New York: Houghton Mifflin.
- Coull, J. T. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging, and psychopharmacology. *Progress in Neurobiology*, *55*(4), 343-361.
- Crawford, J. D. (1991). The relationship between tests of sustained attention and fluid intelligence. *Personality and Individual Differences*, *12*(6), 599-611.
- Deary, I. J. (2008). Why do intelligent people live longer? *Nature*, *456*, 13 November, doi:10.1038/456175a.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13-21.
- Duncan, J. (2003). Intelligence tests predict brain response to demanding task events. *Nature Neuroscience*, *6*(3), 207-208.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428-1446.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475-483.
- 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.
- Flanagan, D. P., & Dixon, S. G. (2014). The Cattell-Horn-Carroll theory of cognitive abilities. *Encyclopedia of Special Education*. 22 Jan 2014. doi: 10.1002/9781118660584.ese0431
- 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*(2), 172-179.
- Frostig, M., Horne, D., & Miller, A. M. (1972). *The developmental program in visual perception*. Chicago: Follett.

- Gottfredson, L. (1997). Why *g* matters: The complexity of everyday life. *Intelligence*, 24, 79-132.
- Gottfredson, L. (2004). Intelligence: Is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? *Journal of Personality and Social Psychology*, 86, 174-199.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Gustafsson, J. E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186-42). New York, NY: Macmillan.
- Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: Inhibition and attentional control. *Neuroimage*, 50(3), 1313-1319.
- Herrnstein, R. J., Nickerson, R. S., Desanchez, M., & Swets, J. A. (1986). Teaching thinking skills. *American Psychologist*, 41(11), 1279-1289.
- Hoffman, M. D., & Gelman, A. (2011). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *arXiv preprint arXiv:1111.4246*.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), 346-363.
- Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, 71, 449-474.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences USA*, 105, 781-791.
- Jaeggi, S. M., Buschkuhl, M., Jonidas, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the USA*, 108, 10081-10086.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- 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(6), 978-990. doi: 10.1111/j.1467-7687.2009.00846.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773-795.
- Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test*. John Wiley & Sons, Inc.
- Klauer, K. J., & Phe, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, 78(1), 85-123. doi: 10.3102/003465430731340
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., & Dahlstrom K. (2005). Computerized training of working memory in children with ADHD – a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44(2), 177-186.
- Kramer, A. F., & Willis, S. L. (2002). Enhancing the cognitive vitality of older adults. *Current Directions in Psychological Science*, 11(5), 173-177.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian modeling for cognitive science: A practical course*. Cambridge, UK: Cambridge University Press.
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1, 6-21.
- McDougall, S., & House, B. (2012). Brain training in older adults: Evidence of transfer to memory span performance and pseudo-Matthew effects. *Aging, Neuropsychology, and Cognition*, 19(1-2), 195-221.
- Mogle, J. A., Lovett, B. J., Strawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory and fluid intelligence. *Psychological Science*, 19(11), 1071-1077.
- Necka, E. (1996). The attentive mind: Intelligence in relation to selective attention, sustained attention, and dual task performance. *Polish Psychological Bulletin*, 27, 3-24.
- Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35, 73-89.
- Posner, M. I. (1978). *Chronometric explorations of mind*. Lawrence Erlbaum.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 12, 25-42.
- R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Raven, J., Raven, J. C., & Court, J. H. (2003, updated 2004) *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, 1(3), 86-89.
- Reinartz, A., & Reinartz, E. (1974). *Visuelle Wahrnehmungsfoerderung*. Hannover: Schroeder.
- Roberts, R. D., Beh, H. C., Spilbury, G., & Stankov, L. (1991). Evidence for an attentional model of human intelligence using the competing task paradigm. *Personality and Individual Differences*, 12(5), 445-455.
- Roberts, R. D., Beh, H. C., & Stankov, L. (1988). Hick's law, competing-task performance, and intelligence. *Intelligence*, 12(2), 111-130.

- Rockstroh, S., & Schweizer, K. (2004). The effect of retest practice on the speed-ability relationship. *European Psychologist, 9*(1), 24-31.
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*(8), 1029-1040.
- Rueda, M. R., Rothbart, R. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences, 102*(41), 14931-14935.
- Ruisel, I. (2003). State of research into intelligence at the beginning of the 21st century – and what next? *Studia Psychologica, 45*(4), 261-283.
- Schweizer, K. (2001). Preattentive processing and cognitive ability. *Intelligence, 29*(2), 169-186.
- Schweizer, S., Hampshire, A., & Dalgleish, T. (2011). Extending brain-training to the affective domain: Increasing cognitive and affective executive control through emotional working memory training. *PLoS One, 6*(9), e24372.
- Schweizer, K., Moosbrugger, H., & Goldhammer, F. (2005). The structure of the relationship between attention and intelligence. *Intelligence, 33*(6), 589-611.
- Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition, 9*, 313-323. doi:10.1006/ccog.2000.0450
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50*(3-4), 3-4.
- Snow, R. E., Kyllonen, C. P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 47-103). Hillsdale, NJ: Erlbaum.
- Stankov, L. (1988). Aging, attention, and intelligence. *Psychology and Aging, 3*(1), 59-74.
- Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology, 49*(3), 607-627.
- Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences, 105*(19), 6791-6792.
- Sturm, W., Orgass, B., & Hartje, W. (2001). *AixTent: A computerised training of four attention functions - a training on alertness, vigilance, selective attention, divided attention*. Bonn: Phoenix Software.
- te Nijenhuis, J., van Vianen, A. E. M., & van der Flier, H. (2007). Score gains on g-loaded tests: No g. *Intelligence, 35*, 283-300.
- Tucha, O., Tucha, L., Kaumann, G., König, S., Lange, K. M., Stasik, D., Streather, Z., Engelschalk, T., & Lange, K. W. (2011). Training of attention functions in children with attention deficit hyperactivity disorder. *Attention Deficit Hyperactivity Disorder, 3*, 271-283.
- Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language, 54*(1), 68-80.
- Wass, S. V., Scerif, G., & Johnson, M. H. (2012). Training attentional control and working memory – Is younger, better? *Developmental Review, 32*(4), 360-387. doi: 10.1016/J.Dr.2012.07.001
- Zimmermann, P., Fimm, B., & Gondan, M. (2002). *Testbatterie zur Aufmerksamkeitsprüfung für Kinder:(KITAP);[das Schloß der Geister]*. Psytest.