Developmental Science

Developmental Science 14:3 (2011), pp 591–601 DOI: 10.1111/j.1467-7687.2010.01022.x

PAPER

Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: a controlled, randomized study

Sissela Bergman Nutley,[1,2] Stina Söderqvist,[1,2] Sara Bryde,[1,2] Lisa B. Thorell,[2,3,4] Keith Humphreys[5] and Torkel Klingberg[1,2]

1. Department of Neuroscience, Karolinska Institutet, Sweden
2. Stockholm Brain Institute, Karolinska Institutet, Sweden
3. Department of Clinical Neuroscience, Karolinska Institutet, Sweden
4. Karolinska Institutet, Center of Neurodevelopmental Disorders, Sweden
5. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden

Abstract

*Fluid intelligence (Gf) predicts performance on a wide range of cognitive activities, and children with impaired Gf often experience academic difficulties. Previous attempts to improve Gf have been hampered by poor control conditions and single outcome measures. It is thus still an open question whether Gf can be improved by training. This study included 4-year-old children (*N = 101*) who performed computerized training (15 min/day for 25 days) of either non-verbal reasoning, working memory, a combination of both, or a placebo version of the combined training. Compared to the placebo group, the non-verbal reasoning training group improved significantly on Gf when analysed as a latent variable of several reasoning tasks. Smaller gains on problem solving tests were seen in the combination training group. The group training working memory improved on measures of working memory, but not on problem solving tests. This study shows that it is possible to improve Gf with training, which could have implications for early interventions in children.*

Introduction

Fluid intelligence (Gf) refers to the ability to, independent of previous knowledge, identify patterns and relations and infer and implement rules (Horn & Cattell, 1966). Gf predicts performance on a wide range of cognitive activities, and low Gf in children is a predictor of academic difficulties (Lynn, Meisenberg, Mikk & Williams, 2007). Gf is impaired in many clinically defined groups, including children with hydrocephalus (Dalen, Bruaroy, Wentzel-Larsen & Laegreid, 2008), post-traumatic brain injury (Ewing-Cobbs, Prasad, Kramer, Cox, Baumgartner, Fletcher, Mendez, Barnes, Zhang & Swank, 2006), Down's syndrome and in children who have undergone brain radiation treatment (Hall, Adami, Trichopoulos, Pedersen, Lagiou, Ekbom, Ingvar, Lundell & Granath, 2004). The question whether Gf can be improved is therefore of great relevance, but the dominant opinion has been that Gf is a fixed trait, unlike crystallized intelligence which is under the influence of learned knowledge (Horn & Cattell, 1966).

Gf training has been attempted since the 1980s with different strategies, tutored problem solving being one of the most common interventions in the older population (Plemons, Willis & Baltes, 1978). This strategy has also been used in children where divergent thinking, inductive reasoning or creative problem solving has been taught through classroom exercises resulting in increases on measures of Gf (Hamers, de Koning & Sijtsma, 1998; Herrnstein, Nickerson, de Sánchez & Swets, 1986; Klauer & Willmes, 2002; Stankov, 1986). None of these reasoning studies, however, have been randomized or had an active control group, making it difficult to assess whether these results were due to the effect of motivation and expectation on performance on IQ tests (Dickstein & Kephart, 1972; Oken, Flegal, Zajdel, Kishiyama, Haas & Peters, 2008). Even though the inclusion of an active control/placebo group is standard procedure in evaluating pharmacological interventions, this is not yet common practice in evaluating cognitive interventions (Klingberg, 2010).

An alternative to tutored problem solving is training of a more basic function potentially underlying reasoning such as working memory (WM) or attention. WM is the ability to hold a limited amount of information in mind in the face of distraction and is required to solve complex problems keeping competing solutions in mind while inferring rules. WM is impaired in most groups with low

Address for correspondence: Sissela Bergman Nutley, Karolinska Institutet, Department of Neuroscience, Retzius Laboratoriet, Stockholm, Sweden; e-mail: sissela.bergman.nutley@ki.se or torkel.klingberg@ki.se

Gf (Levin, Hanten, Chang, Zhang, Schachar, Ewing-Cobbs & Max, 2002), but also in children with Attention Deficit Hyperactivity Disorder (ADHD) (Martinussen, Hayden, Hogg-Johnson & Tannock, 2005; Westerberg, Hirvikoski, Forssberg & Klingberg, 2004) and with learning difficulties (McLean & Hitch, 1999). There is a high correlation between Gf and WM, ranging from 0.4 to 0.9, depending on the specific tasks used and whether a latent variable approach or single task measurements are used (Bühner, Krumm & Pick, 2005; Conway, Kane & Engle, 2003; Engle, Tuholski, Laughlin & Conway, 1999; Kyllonen & Christal, 1990).

In a recent study by Jaeggi and colleagues, Gf improvements were found after training on a dual n-back task when compared to a passive control group (Jaeggi, Buschkuehl, Jonides & Perrig, 2008). However, Gf was assessed with only one single test (Bochumer Matrizen-test) which was administered with a non-standardized procedure (10 min speeded version). The interpretation of this study is thus still open to discussion (Moody, 2009; Sternberg, 2008). Since the definition of Gf itself stems from factor analytical methods, using the shared variance of several tests to define the Gf factor, a similar method should be used to measure gains in Gf. Another issue raised by Sternberg (2008) is that the use of only one single training task makes it difficult to infer if the training effect was due to some specific aspect of the task rather than the general effect of training a construct. However, there are some studies using several WM tasks to train that have also shown transfer effects to reasoning tasks (Klingberg, Fernell, Olesen, Johnson, Gustafsson, Dahlström, Gillberg, Forssberg & Westerberg, 2005; Klingberg, Forssberg & Westerberg, 2002), while other WM training studies have failed to show such transfer (Dahlin, Neely, Larsson, Backman & Nyberg, 2008; Holmes, Gathercole, Place, Dunning, Hilton & Elliott, 2009; Thorell, Lindqvist, Bergman Nutley, Bohlin & Klingberg, 2009). Thus, it is still unclear under which conditions effects of WM training transfer to Gf.

Other intervention studies have included training of attention or executive functions. Rueda and colleagues trained attention in a sample of 4- and 6-year-olds and found significant gains in intelligence (as measured with the Kaufman Brief Intelligence Test) in the 4-year-olds but only a tendency in the group of 6-year-olds (Rueda, Rothbart, McCandliss, Saccomanno & Posner, 2005). A large training study with 11,430 participants revealed practically no transfer after a 6-week intervention (10 min/day, 3 days a week) of a broader range of tasks including reasoning and planning or memory, visuo-spatial skills, mathematics and attention (Owen, Hampshire, Grahn, Stenton, Dajani, Burns, Howard & Ballard, 2010). However, this study lacked control in sample selection and compliance. In summary, it is still an open question to what extent Gf can be improved by targeted training.

The main aims of this study were therefore to investigate: (1) if Gf is improved through computerized training on non-verbal reasoning (NVR) tasks; and (2) if training WM or NVR would result in any transfer to measures of the non-trained construct, Gf and WM, respectively. An unresolved question in cognitive training is whether training should be focused on one construct, or if time would be more efficiently used and lead to larger transfer if divided between several cognitive domains. In order to investigate this we also included a third group, the combined group (CB), training both WM and NVR, to see if combining two types of training would result in any synergistic effects.

For all paradigms, we investigated (1) the training effects on tests similar to the trained tasks (i.e. trained tests); (2) tests assessing the same construct but differing in their composition and presentation (i.e. transfer tests within the same construct); and (3) tests assessing the non-trained construct (i.e. transfer tests between constructs). The NVR training tasks were based on Gf loaded tests from the Leiter Test Battery (Roid & Miller, 1997).

Given the importance of WM and Gf in academic achievement and learning, it would be preferable to train impaired children as early as possible, preventing further missed learning opportunities. Therefore we included preschool children in this study to test the feasibility of training children with as low cognitive level as 4-year-olds. The WM training in this study was visuo-spatial in nature for two reasons: (1) because not all participants had basic knowledge of letters and (2) because Gf has been shown to be more closely related to the visuo-spatial domain than the verbal (Kane, Hambrick, Tuholski, Wilhelm, Payne & Engle, 2004).

Methods

Participants and procedure

The study was designed as a double blinded, randomized, controlled investigation using several training tasks of the same construct and several transfer tests for each construct of interest. Children were randomized (after stratification by gender) to a 5-week computerized intervention of either WM training, NVR training, a combination of WM and NVR training (CB training) or a placebo group (PL training). The active control/placebo group trained on the same tasks as the CB training group but stayed at the lowest difficulty level (using a non-adaptive algorithm) throughout the whole training period.

We included 112 children aged 4–4.5 years (68 boys, mean age = 51.2 months, SD = 3.03) in the study. Subjects were recruited through preschools, flyers, the lab webpage and advertisements in the local newspapers in Stockholm. A prerequisite for participation was access to a PC computer with internet; however, laptops were lent to two subjects with no PC computer available. The group sizes were WM training, n = 24; NVR training,

$n = 25$; CB training, $n = 27$ and PL training, $n = 25$. Before the training period began the participants and their parents/guardians came to the lab where the children went through neuropsychological assessment (with blinded testers) while the parents were coached in the training program by the researchers. The parents were instructed to support and supervise the training. If the children had difficulties handling the mouse (or started playing with the mouse) the parents were instructed to help them and let the children point at the screen instead. After the training conclusion, the neuropsychological assessment was repeated for the children while the parents filled out a questionnaire on motivational experiences during the training. The study was approved by the local ethics committee and informed consent was obtained from the children's guardians.

Materials

Neuropsychological assessment

The problem solving tests used were: Repeated Patterns (RP), Sequential Orders (SO) and Classifications (CL) subtests from the Leiter battery (Roid & Miller, 1997), Raven's Coloured Progressive Matrices (CPM) (Raven, 1998) and Block Design from WPPSI (Wechsler, 2004). In the RP, SO and CL tests, subjects are required to choose cards or shapes in different colours and sizes to fill in the gaps in order to either repeat the pattern displayed, continue the sequence or classify/match the cards or shapes based on the features of the shape (i.e. size and colour). In Raven's CPM, participants are presented with an incomplete matrix and have to select one of six choices to complete it. There are three sets (A, AB and B) of 12 items each. Set A contains items with a figure that require the subject to complete the missing parts (gestalt continuation). Sets AB and B contain items where the matrices gradually shift from a coherent whole (gestalt) to four separate symbols creating increasingly more difficult items that require identification of relations between the separate symbols within the matrix as well as comparisons between the relations to the response alternatives (Raven, 1998). In Block Design, subjects are required to as quickly as possible reproduce a pattern shown on a card using red and white blocks. The completion of each item is timed and the faster it is correctly completed the higher the score achieved. The Raven's CPM and Block Design are transfer tests *within* construct for the NVR and CB groups and transfer tests *between* constructs for the WM training group.

The tests used to assess the memory domain were: a visuo-spatial grid task (Bergman Nutley, Söderqvist, Bryde, Humphreys & Klingberg, 2010; Westerberg *et al.*, 2004), the Odd One Out from the Automated Working Memory Assessment (AWMA; Alloway, 2007) and the Word Span test (Thorell & Wåhlstedt, 2006). The visuo-spatial grid test consists of a 4 by 4 matrix in which yellow circles are presented sequentially for the subject to relocate in the correct order by pointing to the screen at the end of each sequence.

In Odd One Out, the participant views three simultaneously presented framed shapes and has to identify the odd one out by pointing to it. The task is then to keep its location in memory while viewing a new set of shapes and again identifying the odd one out. When cued with three empty frames, the subject has to point to where the odd shapes were in the order they were displayed. The Odd One Out is a span test and the number of sets of shapes displayed begins with only one set of shapes and increases with performance.

The Word Span test is a verbal short-term memory (STM) test similar to the digit span forward subtest from WISC-III (Wechsler, 1991) but with unrelated nouns instead of numbers. All three memory tests were span tests increasing in load for every correctly passed level. The Odd One Out and Word Span tests are transfer tests *within* constructs for the WM and CB groups and *between* constructs for the NVR group.

The training program

The study included three different training paradigms: WM training, NVR training and CB training. The WM training was the same as described in Thorell *et al.* (2009) developed by Cogmed Systems Inc. There were seven different versions of visuo-spatial WM tasks, out of which three were trained every day on a rotating schedule. Briefly, the tasks all consisted of a number of animated figures presented in different settings (e.g. swimming in a pool, riding on a rollercoaster). Some of the figures (starting with two figures and then increasing in number depending on the child's performance) made a sound and changed colour during a short time period. The task then consisted of remembering which figures had changed colour and in what order this had occurred. The NVR training paradigm was developed for this study based on the three Gf loaded tests (i.e. RP, SO and CL) from the Leiter Battery (Roid & Miller, 1997). The tasks were modified so that new problems could be automatically generated by a computer algorithm, allowing us to present a wide range of reasoning problems. All items consisted of geometrical figures.

Repeated patterns (RP). The task was to identify the pattern presented and select from answer cards the correct ones that complete the pattern. The top of the screen had a number of cards showing shapes and the two last slots were empty. At the bottom of the screen there were answer cards that the child could choose from, click on and drag to the empty slots. The difficulty was altered by changing the number of answer cards, the number of cards in the top row, number of rows (one or two), the length of the pattern, and by the number of modalities that would change (size, colour, shape, filled slice position and number of dots on the shape) (see Figure 1A).

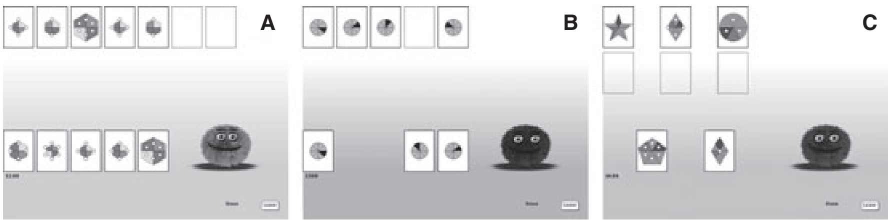

Figure 1 *Examples from the three tasks in the NVR training program consisting of Repeated Patterns (A), Sequential Orders (B) and Classifications (C). Repeated Patterns requires the child to identify the pattern and select the correct two cards from the bottom to continue the pattern; Sequential Orders contains a gap in the progression for the child to fill out by choosing the correct card from the bottom row; and Classifications requires matching of the two bottom cards with the two correct ones in the top row based on one specific rule.*

Sequential order (SO). The task was to understand the relationship (logical progression) between different stimuli and select from answer cards the correct ones to complete the sequence by placing them in the beginning, middle or end of the row of cards. The presentation was similar to that of RP. The difficulty was adjusted by altering the number of answer cards, the number of cards in the top row, number of empty slots and number of modalities that would change (size, colour, shape, filled slice positions, number of dots on a shape and position) (see Figure 1B).

Classification (CL). This task required the child to sort answer cards so that they matched the presented cards on the top row on a specific modality, e.g. colour. The top row consisted of a number of cards with different shapes, colours, positions of filled slices, number of dots on the shapes and sizes, and the child had to match the two answer cards to two of the top row cards. The difficulty was altered by changing the number of cards on the top row, the number of modalities differing among the cards in an item, the number of possible matches for one of the answer cards (one or two) (see Figure 1C).

For all three tasks, the child made a response by clicking with the mouse on a card and dragging it to an empty slot and then pressing 'enter'. If the answer was correct, the child would see a star before the next item was presented. If the answer was incorrect, the computer would move the answer cards to the correct slots. The parents were instructed to have only an encouraging role until the child made an incorrect response, and when the computer showed the correct answer they should explain it.

The hierarchy of levels was carefully piloted on an independent sample of 28 children (between the ages of 4 and 6). The NVR training group trained all three NVR tasks every day. The CB training group trained two NVR tasks and two WM tasks every day on a rotating schedule. All three paradigms were adaptive, indicating that the level of difficulty was adjusted based on each child's performance. The PL training had the same paradigm as the CB training, the only difference being that they remained at the easiest level throughout the training. All training was performed in identical ways, at home with the parents and lasting around 15 minutes/day, 5 days/week for 5–7 weeks, until 25 sessions had been performed. Performance was automatically recorded and uploaded to a server, which allowed us to confirm compliance. Parents received feedback by email or phone once a week.

Results

Training and baseline performance

One-way analysis of variance (ANOVA) using SPSS version 16.0.1 with the fixed factor group (i.e. NVR, WM, CB and PL training) and performance at baseline (T1) as the dependent variable did not show any significant difference between the groups for any of the tests (all $Fs(3, 101) < 2.14$, $p > .05$). Furthermore, the four groups did not differ in terms of age, age at preschool admission, hours of computer games played prior to the study entry or education level of the parents (all $Fs(3, 101) < 1.62$, $p > .05$ and $Hs(3, 100) < 4.55$, $p > .05$ for education level as tested with a Kruskal-Wallis test). A total of 101 out of 112 children tested at T1 completed at least 20 out of the 25 sessions of training ($M = 24.09$, $SD = 1.61$) and were evaluated at training completion (T2). The failures to complete at least 20 sessions were due to computer problems ($n = 3$), illness or time restraints ($n = 6$) and motivational problems ($n = 2$). No significant difference in drop-out between the groups was observed.

The training time per day was approximately 15 minutes and the estimated mean for each group was 17.1, 14.6, 16.0 and 16.1 minutes for the NVR, WM, CB and PL training groups, respectively. After training, the parents rated congruency (scale 1 to 5) to statements regarding how motivating, challenging, unvarying and fun they thought their children experienced the training and there were no significant group differences (Kruskal-Wallis test, all $Hs(3, 101) < 7.82$, $p > .05$).

Effects on problem solving tests

The four transfer measurements (i.e. set A, AB and B on the Raven's CPM and Block Design) were highly correlated with each other (Table 1), so in order to assess Gf

Table 1 *Bivariate correlations for all tests at T1*

Tests	1	2	3	4	5	6	7	8
WM/STM tests								
1. Grid Task	1							
2. Odd One Out	0.45**	1						
3. Word Span	0.19	0.22*	1					
Problem solving tests								
4. Leiter	0.27**	0.17	0.10	1				
5. Raven set A	0.37**	0.13	0.19	0.24*	1			
6. Raven set AB	0.34**	0.11	0.18	0.30**	0.42**	1		
7. Raven set B	0.34**	0.40**	0.18	0.30**	0.38**	0.50**	1	
8. Block Design	0.55**	0.46**	0.21*	0.30**	0.51**	0.42**	0.56**	1

* $p < .05$; ** $p < .01$.

gains as a result of training, we analysed the construct as a latent factor. For each time point (separately) the scores of the four tests were modelled as normally distributed variables with means specified to be independent linear functions of a continuous latent variable. Maximum likelihood estimation was used to fit the models, using Amos for Windows (Version 16.0.1). The single latent variable model provided a good description of the covariance structure at both time points; at T1 the Akaike Information Criterion (AIC) value for the model was only slightly higher than that of the 'saturated' (fully specified covariance structure) model (difference = 1.2) and at T2 the AIC value was smaller for the latent variable model than for the saturated model (difference = −2.7). At T1 the latent factor analysis revealed similar loadings from the four subtests (standardized regression weights 0.62, 0.63, and 0.72 from the Raven's CPM A, AB and B subtests, respectively, and 0.76 from Block Design). The standardized regression weights at T2 were 0.46, 0.77, and 0.76 from the Raven's CPM A, AB and B subtests, respectively, and 0.77 for Block Design. The expected value of the latent variable from T2 (given the tests scores) was used as a dependent variable in an ANCOVA with group as fixed factor, age in months and the expected value of the latent variable at T1 as covariates. In the event of a significant or marginally significant ($p < .10$) group effect, planned comparisons were performed to investigate group differences between each one of the three training groups and the placebo group (unpaired *t*-tests, two tailed). There was a significant training effect on the group level ($F(3, 101) = 4.64$, $p = .005$) and planned comparisons revealed that the NVR training group ($p = .02$) and the CB training group ($p = .05$) had improved significantly more than the PL training group (Table 2).

To further assess the specificity of training effects on individual measures, similar ANCOVAs were performed with individual task performance at T2 as the dependent variable, group as the fixed factor, age in months and performance at T1 as covariates (see Table 2 for results). In the event of a significant or marginally significant ($p < .10$) group effect, planned comparisons were performed to investigate group differences between each one of the three training groups and the placebo group (unpaired *t*-tests, two tailed). Means at pre- and post-training are shown in Table 2 as well as standardized change for each group ($(M_{T2} - M_{T1})/SD_{pooled\ T1}$). The measures of standardized change were used to illustrate the gain in trained groups as well as the placebo group, as it was expected that all groups would show some level of increase (Figure 2). Subtracting the change of the PL training group from each of the training groups thus gives the effect sizes (Cohen's delta) for each training group respectively.

The NVR and CB training groups improved significantly on the test similar to the trained one, the Leiter total score (consisting of the sum of RP, SO and CL). This test was included as a confirmation that the training itself worked, being similar to the trained tasks but differing in administration, and should not be viewed as an objective test of cognitive improvement.

The training did not show effects on the total score of Raven's CPM but since the scores between sets seemed to differ, they were split (in accordance with the manual recommendation) into sets A, AB and B. There was then a significant effect for set B for the NVR training group. There was also a significant training effect on the Block Design test for the NVR training group.

Effects on WM/STM tests

Due to the different nature of the two transfer tests in the memory domain (Odd One Out and Word Span), no latent variable analysis was performed for these. The WM and CB training groups both improved significantly on the trained test of visuo-spatial WM (the Grid task) (Table 2). This test was included as confirmation that the training itself worked, being similar to the trained tasks but differing in administration, and should not be viewed as an objective test of cognitive improvement. The training also showed effects on the memory score of the transfer test, the Odd One Out for the WM and CB groups with a tendency for the NVR training group, indicating transfer *within* construct for the WM and CB training groups and *between* constructs for the NVR training group. Effects on the Word Span test did not reach significance at the group level. However, because improvement on this test was shown in a previous WM

Table 2 *Means* (M) *and standard deviations* (SD) *for each group pre- (T1) and post- (T2) training. Standardized changes* ((M_{T2}–M_{T1})/$SD_{pooled\ T1}$) *are also listed and* p*-values refer to comparisons with the placebo group. The overall group effects from the ANCOVAs are listed in the last column*

Tests	Training group	T1 *M* (*SD*)	T2 *M* (*SD*)	SC $_{T2–T1}$	ANCOVA
WM/STM tests					
Grid Task	NVR	20.3 (10.1)	25.5 (11.6)	0.39	
	WM	24.4 (18.0)	50.6 (24.8)	1.95 ***	$F(3, 90) = 18.99$, $p < .001$
	CB	18.1 (12.6)	46.2 (20.1)	2.10 ***	
	PL	20.0 (12.4)	21.8 (7.5)	0.13	
Odd One Out	NVR	7.7 (2.6)	9.4 (3.8)	0.67 +	
	WM	6.9 (3.2)	9.7 (4.3)	1.12 *	$F(3, 98) = 2.40$, $p = .07$
	CB	7.8 (2.3)	9.9 (3.7)	0.84 *	
	PL	7.3 (2.0)	7.7 (2.7)	0.17	
Word Span	NVR	14.2 (6.3)	15.3 (4.8)	0.20	
	WM	14.2 (6.5)	17.0 (6.0)	0.49	$F(3, 99) = 1.22$, $p = 0.3$
	CB	13.2 (5.1)	14.6 (5.2)	0.24	
	PL	12.6 (5.3)	14.1 (5.3)	0.26	
Problem solving tests					
Leiter	NVR	26.0 (6.0)	34.1 (10.6)	1.21 *	
	WM	26.2 (7.9)	29.3 (6.5)	0.47	$F(3, 98) = 4.63$, $p = .005$
	CB	25.2 (7.9)	35.1 (11.7)	1.48 *	
	PL	26.4 (5.2)	30.3 (9.2)	0.57	
Raven Set A	NVR	7.7 (1.5)	7.7 (1.3)	0.02	
	WM	6.9 (1.8)	7.6 (1.1)	0.39	$F(3,101) = 0.01$, $p = .99$
	CB	6.8 (1.4)	7.6 (1.3)	0.42	
	PL	6.5 (2.0)	7.4 (1.4)	0.54	
Raven Set AB	NVR	3.1 (2.2)	4.8 (2.6)	0.90	
	WM	2.8 (1.5)	3.9 (2.2)	0.55	$F(3, 99) = 1.78$, $p = .16$
	CB	2.8 (1.9)	5.0 (2.6)	1.12	
	PL	2.6 (2.2)	3.9 (2.5)	0.67	
Raven Set B	NVR	2.7 (2.3)	4.3 (2.1)	0.92 *	
	WM	2.9 (1.4)	2.6 (1.7)	−0.18	$F(3, 97) = 4.56$, $p = .001$
	CB	2.1 (1.1)	3.5 (2.0)	0.84	
	PL	2.6 (1.4)	3.1 (1.8)	0.29	
Block Design	NVR	21.8 (3.6)	24.5 (4.0)	0.76 ***	
	WM	21.8 (4.2)	22.9 (3.2)	0.33	$F(3, 101) = 4.15$, $p = .008$
	CB	21.7 (3.8)	22.7 (3.4)	0.29	
	PL	21.2 (2.9)	21.9 (3.5)	0.19	
Gf latent variable	NVR	0.4 (2.8)	0.9 (2.7)	0.18 *	
	WM	0.2 (2.4)	−0.5 (2.2)	−0.30	$F(3, 101) = 4.64$, $p = .005$
	CB	−0.3 (2.1)	0.2 (2.4)	0.34 *	
	PL	−0.3 (2.4)	−0.6 (2.3)	−0.13	

+ $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

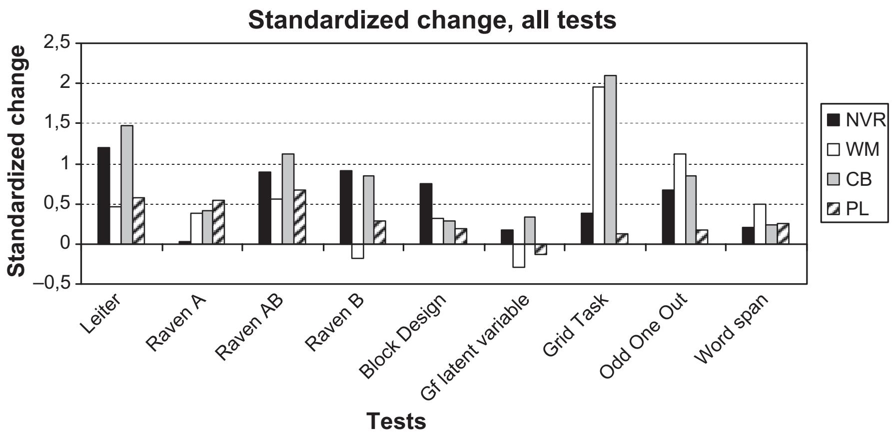

Figure 2 *Standardized change* ((M_{T2}–M_{T1})/$SD_{pooled\ T1}$) *for each of the groups for all of the tests.*

training study in preschool children (Thorell *et al*., 2009), we performed a guided test comparing the WM to the PL group, which showed a marginally significant effect (p = .05, one-tailed t-test).

Interaction effects of combined training

In order to investigate possible interactive effects of training both WM and NVR, we analysed the effect of

Table 3 *Regression coefficients for the interaction analysis. The table contains the standardized coefficients (β) from the linear regression analysis using factors instead of groups. The data were modelled with performance at T2 as the dependent variable, age, performance at T1 and groups coded in accordance with the amount of training received on the factors WM, NVR or WM × NVR as the independent variables (e.g. coding for CB group was 1,1,1 and coding for NVR group was 0,2,0)*

Tests	R^2 (model)	WM factor (β)	NVR factor (β)	WM × NVR factor (β)
WM/STM tests				
Grid Task	0.52***	0.50***	0.082	0.22**
Odd One Out	0.32***	0.25*	0.17+	0.0042
Word Span	0.41***	0.15	0.033	−0.11
Problem solving tests				
Leiter	0.44***	−0.048	0.18*	0.18*
Raven set A	0.12*	0.012	−0.0064	0.0070
Raven set AB	0.38***	−0.038	0.11	0.13
Raven set B	0.29***	−0.14	0.23*	0.082
Block design	0.54***	0.067	0.27***	−0.11
Gf	0.68***	−0.056	0.16*	0.088

+ $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

training not coded as four different groups, but coded depending on the amount of NVR or WM trained. We labelled each condition based on the amount of training on NVR or WM, and also included an interaction term. Thus, the coding for the three training factors (WM, NVR and WM × NVR) for each group was: for WM training group 2, 0, 0; for NVR training group, 0, 2, 0; for CB training group, 1, 1, 1; and for PL training group, 0, 0, 0. This controlled for the fact that the CB group only trained each construct half of the time compared to the single construct groups (WM or NVR training). We performed a linear regression with T2 performance as the dependent variable while T1 performance, the three training factors and age were set as independent variables. There was a significant interaction effect (WM × NVR) for two of the tests, the Leiter total score (R^2 = 0.43, $p < .001$, for the model and $\beta = 0.18$, $p = .03$ for the interaction term) and the Grid task ($R^2 = 0.52$, $p < .001$, for the model and $\beta = 0.22$, $p = .008$, for the interaction term), indicating interaction for outcome measures of trained tests only, but no interaction effect on transfer tests within or between constructs. This set of analyses also confirmed the previously reported main effects with the WM factor showing effects on the Grid Task and Odd One Out, and NVR factor showing effects on Leiter, Gf as latent variable, Raven's CPM set B, Block Design and a tendency on the Odd One Out (see regression coefficients and p-values in Table 3).

Discussion

This study aimed to investigate whether it is possible to improve Gf by training on either reasoning problems, WM tasks or a combination of both. Gf was improved by the reasoning training but not by the WM training, as shown by analysing the performance on the Gf tasks as a latent variable.

It is unlikely that the improved performance in the active training groups was due to expectancy or motivation since we included a placebo group. In addition, the parents of the children in the placebo group rated the training as equally fun, motivating, and challenging as the parents of children in the other groups. Thus, this condition most likely remained blind to group randomization throughout the study.

Increases in Gf

Training excessively on any given task can lead to improvements on that specific task for two reasons. The first reason is task-specific improvements, such as development of efficient strategies and stimulus-specific priming, and the second reason being enhancement of underlying ability, which would generalize to other tasks relying on the same ability. Since we are only interested in the latter we wanted to evaluate our subjects on tests that relied on the same underlying ability as the trained ones but differed in design, presentation and response mode.

Two tests were used to estimate transfer to Gf: Raven's CPM and Block Design. Similarly to the trained tasks, the Raven's CPM demands making a selection from several alternatives, but unlike the trained tasks, has a fixed number of missing pieces (one) and response alternatives. Raven's CPM contains three sets, A, AB and B, which tap slightly different aspects of reasoning (Raven, 1998; Strauss, Sherman & Spreen, 2006). Set A is constructed to test gestalt continuation, requiring the subjects to complete a figure. Set AB continues with gestalt continuation and proceeds to test analogical reasoning, where the subjects are required to perceive the relation between the now four separate figures and their relation to the response alternatives in order to deduce the logical rule, apply it and select the correct response. Set B consists almost entirely of items testing analogical reasoning.

The effect of NVR training differed between the three sets, with no effect on set A, a stronger but not significant effect on set AB (effect size 0.23) and a significant effect on set B (effect size 0.63). None of the three training tasks in the NVR program contained items requiring completion of a figure, which is consistent with the lack of improvement on set A. Instead, the focus of the training lay in pattern identification, deduction of rules and resisting distracting competing choices, all of which are required for analogical reasoning. NVR training also significantly improved performance of the Block Design test. This test has building blocks, a requirement to replicate a presented pattern and a timed component. Thus it differs from the trained tasks both in design, presentation and response mode. These results demonstrate transfer beyond task-specific improvement. The NVR training group showed transfer both when this was estimated with single tests, as well as when Gf was measured as a latent variable. The magnitude of this improvement was approximately 8% (compared to the placebo group) which is comparable with previously reported gains of Gf of 5–13.5% (Hamers *et al*., 1998; Jaeggi *et al*., 2008; Klauer & Willmes, 2002; Stankov, 1986).

Transfer between constructs

Our results replicate previous findings that it is possible to train WM, and that it transfers to non-trained WM tests (Holmes *et al*., 2009; Klingberg *et al*., 2005; Klingberg *et al*., 2002; Thorell *et al*., 2009). The transfer to these non-trained tests shows that the effect is not simply an improved strategy, but enhancement of an underlying ability.

However, no transfer was seen to reasoning tests, which is in line with some previous studies (Holmes *et al*., 2009; Thorell *et al*., 2009) but not others (Jaeggi *et al*., 2008; Klingberg *et al*., 2005; Klingberg *et al*., 2002). This could mean that WM is not a limiting factor for 4-year-old children solving reasoning problems such as Raven's CPM and Block Design. The moderate correlations between the Grid Task and the reasoning tests (between 0.3 and 0.6, see Table 1) point to the somewhat counterintuitive conclusion that correlation between two underlying abilities is not a sufficient predictor to determine amount of transfer of training effects between these abilities. A similar conclusion was drawn after the lack of training effects on WM after training inhibitory functions (Thorell *et al*., 2009). In that study WM capacity correlated with performance on the inhibitory tasks at baseline ($R = 0.3$). An imaging study also showed that performance on a WM grid task and inhibitory tasks activate overlapping parts of the cortex (McNab, Leroux, Strand, Thorell, Bergman & Klingberg, 2008). Inhibitory training improved performance on the trained tasks, yet there was no transfer seen on WM tasks. The principles governing the type of cognitive training that will transfer are still unclear and pose an important question for future studies.

One way to find these principles may be through understanding the neural mechanisms of training. For example, WM training in 4-year-olds might have a more pronounced effect on the parietal lobe, compared to the less mature frontal lobe. If the transfer to Gf is dependent on prefrontal functions, it may explain the lack of transfer from WM training to Gf in 4-year-olds. In other words, transfer effects may differ with the progression of development.

The single transfer effect seen *between* constructs was for the NVR training group showing a tendency to improvement on the Odd One Out test. This indicates that the NVR training encompasses aspects of WM that are required in this test but not in the Grid Task and Word Span tests.

Training broadly or intensely – principles of cognitive training

Another question that was addressed in this study was whether training both WM and NVR would result in any synergistic effects such as further transfer or higher level of performance on either one or both abilities after training. When training effect was analysed in terms of WM training and NVR training load, the only interaction between the two constructs (WM × NVR) was for the Grid Task and the Leiter tests (Table 3). Both these tests are similar to trained tasks, and the interaction effect indicates that familiarity is not linearly related to amount of training time, but was larger than expected given the difference in amount of training between groups. This set of regression analyses also confirmed our results from the ANCOVA regarding the effects of training WM or NVR. These analyses took into account that the groups differed in the amount of training received, full dose for NVR or WM groups or half dose for the CB group (Table 3). Even though the pattern is not consistent across all tests (see Figure 2), this is interpreted as confirmation of the linear dose effect that was expected to be seen. Our results suggest that the amount of transfer to non-trained tasks within the trained construct was roughly proportionate to the amount of training on that construct. A similar finding, with transfer proportional to amount of training, was reported by Jaeggi *et al*. (2008). This has possible implications for the design of future cognitive training paradigms and suggests that the training should be intensive enough to lead to significant transfer and that training more than one construct does not entail any advantages in itself. The training effect presumably reaches asymptote, but where this occurs is for future studies to determine. It is probably important to ensure that participants spend enough time on each task in order to see clinically significant transfer, which may be difficult when increasing the number of tasks being trained. This may be one of the explanations for the lack of transfer seen in the Owen *et al*. study (2010) (training six tasks in 10 minutes).

Possible neurobiological mechanisms for transfer

Animal research has shown how repeated perceptual (Recanzone, Merzenich, Jenkins, Grajski & Dinse, 1992) and motor training (Nudo, Milliken, Jenkins & Merzenich, 1996) can change synaptic connectivity and strength. Simulation studies suggest that synaptic strength in the networks keeping information online in WM affects the overall capacity (Edin, Klingberg, Johansson, McNab, Tegner & Compte, 2009; Edin, Macoveanu, Olesen, Tegner & Klingberg, 2007). One explanation of the neural underpinnings to training is that capacity is affected through strengthening of synaptic connectivity. Some previous studies on WM training have shown an increase in activation in areas related to WM performance, i.e. prefrontal and parietal cortices (Hempel, Giesel, Garcia Caraballo, Amann, Meyer, Wüstenberg, Essig & Schröder, 2004; Jolles, Grol, Van Buchem, Rombouts & Crone, 2010; Moore, Cohen & Ranganath, 2006; Olesen, Westerberg & Klingberg, 2004), while others have shown reduced activity (Dahlin, Backman, Neely & Nyberg, 2009; for a review see Klingberg, 2010). Dopamine has also been implicated as a possible modulator in a study where WM training led to an alteration in D1 receptor density (McNab, Varrone, Farde, Jucaite, Bystritsky, Forssberg & Klingberg, 2009). It has also been shown that WM training can increase myelination in WM-related brain areas (Takeuchi, Sekiguchi, Taki, Yokoyama, Yomogida, Komuro, Yamanouchi, Suzuki & Kawashima, 2010).

The neural mechanisms underlying NVR training could be similar to those of WM training, although the mechanisms at single neural levels are less explored. To our knowledge, there have not been any neuroimaging studies training NVR extensively. Shorter practice sessions with geometrical analogies have resulted in decreased activity in a bilateral fronto-parietal network (Wartenburger, Heekeren, Preusse, Kramer & van der Meer, 2009). However, the training was too short to generalize and the decreases in brain activity could reflect familiarity effects. Developmental neuroimaging studies suggest a positive correlation with prefrontal activity and reasoning performance (rostrolateral prefrontal cortex in particular) (Crone, 2009; Wright, Matlen, Baym, Ferrer & Bunge, 2008). The transfer seen between the constructs (for the Odd One Out) in this study could possibly be mediated by repeated activation of overlapping neural networks (fronto-parietal/frontal), but needs to be studied.

Training 4-year-olds

The clinical reason for training 4-year-olds was to study how early these interventions could be practically possible if clinically requested. This study shows that it is possible to train both Gf and WM at the developmental age of 4 and it is for future studies to examine training responses in clinical populations.

The optimal ages for cognitive training are as yet unknown. It is often assumed that younger brains are more plastic. One study has suggested that starting piano training at an early age has a larger effect on motor tracts of the brain (Bengtsson, Nagy, Skare, Forsman, Forssberg & Ullen, 2005). However, there is still very little data to provide information of any susceptibility differences between ages, much less evidence of particularly sensitive periods, or windows of opportunity, where cognitive training would have a larger effect. Even though a younger brain might be more plastic, it may also mean that children with immature prefrontal cortex benefit less from training of cognitive tasks, as discussed by Thompson-Schill, Ramscar and Chrysikou (2009). A prerequisite for attaining training effects is the presence of some rudimentary form of the function of interest. It is not worth training ice skating before one is able to walk. Judging from the results of this study, WM and reasoning ability are established enough to be able to train at this age. This is in line with the results from Thorell *et al*. (2009) showing improvements in WM after training in 4–5-year-olds and the results from Rueda *et al*. (2005) that showed an increase in intelligence in 4-year-olds.

Conclusions

In summary, we found that Gf can be improved through 5 weeks of NVR training in 4-year-olds. This type of training might be useful for children with poor intelligence. Early detection and intervention of children who would benefit from NVR and or WM training could possibly prevent falling behind at school and allow learning opportunities that may otherwise be lost due to impaired cognitive capacities. Future studies could investigate the effects of NVR training on a clinical sample, investigate transfer to real-life situations (e.g. academic performance) and assess the long-term effects of this type of training.

Acknowledgements

This work was supported by the Swedish Foundation for Strategic Research, Swedish Research Council, VINNOVA (the Swedish Governmental Agency for Innovation Systems), and the Knut and Alice Wallenberg Foundation. We would like to thank Jonas Beckeman for programming the NVR tasks.

References

Alloway, T.P. (2007). *Automated working memory assessment*. London: Psychological Corporation.

Bengtsson, S.L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, **8** (9), 1148–1150.

Bergman Nutley, S., Söderqvist, S., Bryde, S., Humphreys, K., & Klingberg, T. (2010). Measuring working memory capacity with greater precision in the lower capacity ranges. *Developmental Neuropsychology*, **35** (1), 81–95.

Bühner, M., Krumm, S., & Pick, M. (2005). Reasoning = working memory ≠ attention. *Intelligence*, **33**, 251–272.

Conway, A.R., Kane, M.J., & Engle, R.W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, **7** (12), 547–552.

Crone, E.A. (2009). Executive functions in adolescence: inferences from brain and behavior. *Developmental Science*, **12** (6), 825–830.

Dahlin, E., Backman, L., Neely, A.S., & Nyberg, L. (2009). Training of the executive component of working memory: subcortical areas mediate transfer effects. *Restorative Neurology and Neuroscience*, **27** (5), 405–419.

Dahlin, E., Neely, A.S., Larsson, A., Backman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, **320** (5882), 1510–1512.

Dalen, K., Bruaroy, S., Wentzel-Larsen, T., & Laegreid, L.M. (2008). Intelligence in children with hydrocephalus, aged 4–15 years: a population-based, controlled study. *Neuropediatrics*, **39** (3), 146–150.

Dickstein, L.S., & Kephart, J.L. (1972). Effect of explicit examiner expectancy upon WAIS performance. *Psychological Reports*, **30** (1), 207–212.

Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegner, J., & Compte, A. (2009). Mechanism for top-down control of working memory capacity. *Proceedings of the National Academy of Sciences, USA*, **106** (16), 6802–6807.

Edin, F., Macoveanu, J., Olesen, P., Tegner, J., & Klingberg, T. (2007). Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *Journal of Cognitive Neuroscience*, **19** (5), 750–760.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology, General*, **128** (3), 309–331.

Ewing-Cobbs, L., Prasad, M.R., Kramer, L., Cox, C.S., Jr., Baumgartner, J., Fletcher, S., Mendez, D., Barnes, M., Zhang, X., & Swank, P. (2006). Late intellectual and academic outcomes following traumatic brain injury sustained during early childhood. *Journal of Neurosurgery*, **105** (4 Suppl.), 287–296.

Hall, P., Adami, H.O., Trichopoulos, D., Pedersen, N.L., Lagiou, P., Ekbom, A., Ingvar, M., Lundell, M., & Granath, F. (2004). Effect of low doses of ionising radiation in infancy on cognitive function in adulthood: Swedish population based cohort study. *British Medical Journal*, **328** (7430), 19–24.

Hamers, J.H.M., de Koning, E., & Sijtsma, K. (1998). Inductive reasoning in third grade: intervention promises and constraints. *Contemporary Educational Psychology*, **23**, 132–148.

Hempel, A., Giesel, F.L., Garcia Caraballo, N.M., Amann, M., Meyer, H., Wüstenberg, T., Essig, M., & Schröder, J. (2004). Plasticity of cortical activation related to working memory during training. *American Journal of Psychiatry*, **161** (4), 745–747.

Herrnstein, R.J., Nickerson, R.S., de Sánchez, M., & Swets, J.A. (1986). Teaching thinking skills. *American Psychologist*, **41** (11), 1279–1289.

Holmes, J., Gathercole, S.E., Place, M., Dunning, D.L., Hilton, K.A., & Elliott, J.G. (2009). Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology*, **24** (6), 827–836.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, **57** (5), 253–270.

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, USA*, **105** (19), 6829–6833.

Jolles, D.D., Grol, M.J., Van Buchem, M.A., Rombouts, S.A., & Crone, E.A. (2010). Practice effects in the brain: changes in cerebral activation after working memory practice depend on task demands. *NeuroImage*, **52** (2), 658–668.

Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., & Engle, R.W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, **133** (2), 189–217.

Klauer, K.J., & Willmes, K. (2002). Inducing inductive reasoning: does it transfer to fluid intelligence? *Contemporary Educational Psychology*, **27**, 1–25.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, **14** (7), 317–324.

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

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

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity. *Intelligence*, **14** (4), 389–433.

Levin, H.S., Hanten, G., Chang, C.C., Zhang, L., Schachar, R., Ewing-Cobbs, L., & Max, J.E. (2002). Working memory after traumatic brain injury in children. *Annals of Neurology*, **52** (1), 82–88.

Lynn, R., Meisenberg, G., Mikk, J., & Williams, A. (2007). National IQs predict differences in scholastic achievement in 67 countries. *Journal of Biosocial Science*, **39** (6), 861–874.

McLean, J.F., & Hitch, G.J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, **74** (3), 240–260.

McNab, F., Leroux, G., Strand, F., Thorell, L., Bergman, S., & Klingberg, T. (2008). Common and unique components of inhibition and working memory: an fMRI, within-subjects investigation. *Neuropsychologia*, **46** (11), 2668–2682.

McNab, F., Varrone, A., Farde, L., Jucaite, A., Bystritsky, P., Forssberg, H., & Klingberg, T. (2009). Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science*, **323** (5915), 800–802.

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, **44** (4), 377–384.

Moody, D.E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence*, **37** (4), 327–328.

Moore, C.D., Cohen, M.X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *Journal of Neuroscience*, **26** (43), 11187–11196.

Nudo, R.J., Milliken, G.W., Jenkins, W.M., & Merzenich, M.M. (1996). Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *Journal of Neuroscience*, **16** (2), 785–807.

Oken, B.S., Flegal, K., Zajdel, D., Kishiyama, S., Haas, M., & Peters, D. (2008). Expectancy effect: impact of pill administration on cognitive performance in healthy seniors. *Journal of Clinical and Experimental Neuropsychology*, **30** (1), 7–17.

Olesen, P.J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, **7** (1), 75–79.

Owen, A.M., Hampshire, A., Grahn, J.A., Stenton, R., Dajani, S., Burns, A.S., Howard, R.J., & Ballard, C.G. (2010). Putting brain training to the test. *Nature*, **465** (7299), 775–778.

Plemons, J.K., Willis, S.L., & Baltes, P.B. (1978). Modifiability of fluid intelligence in aging: a short-term longitudinal training approach. *Journal of Gerontology*, **33** (2), 224–231.

Raven, J.C. (1998). *Manual for Raven's Progressive Matrices*. Oxford: Oxford Psychologists Press.

Recanzone, G.H., Merzenich, M.M., Jenkins, W.M., Grajski, K.A., & Dinse, H.R. (1992). Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *Journal of Neurophysiology*, **67** (5), 1031–1056.

Roid, G.H., & Miller, L.J. (1997). *Leiter International Performance Scale-Revised: Examiner's Manual*. Wood Dale, IL: Stoelting Co.

Rueda, M.R., Rothbart, M.K., McCandliss, B.D., & Posner, P. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences, USA*, **102**, 14931–14936.

Stankov, L. (1986). Kvashchev's experiment: can we boost intelligence? *Intelligence*, **10**, 209–230.

Sternberg, R.J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences, USA*, **105** (19), 6791–6792.

Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropscyhological tests: Adminstration, norms and commentary* (3rd edn.). New York: Oxford University Press.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., Yamanouchi, T., Suzuki, S., & Kawashima, R. (2010). Training of working memory impacts structural connectivity. *Journal of Neuroscience*, **30** (9), 3297–3303.

Thompson-Schill, S.L., Ramscar, M., & Chrysikou, E.G. (2009). Cognition without control: when a little frontal lobe goes a long way. *Current Directions in Psychological Science*, **18** (5), 259–263.

Thorell, L.B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, **12** (1), 106–113.

Thorell, L.B., & Wåhlstedt, C. (2006). Executive functioning deficits in relation to symptoms of ADHD and/or ODD in preschool children. *Infant and Child Development*, **15**, 503–518.

Wartenburger, I., Heekeren, H.R., Preusse, F., Kramer, J., & van der Meer, E. (2009). Cerebral correlates of analogical processing and their modulation by training. *NeuroImage*, **48** (1), 291–302.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* - third edition. (Psykologiförlaget AB, Stockholm, trans.). New York: Psychological Corporation.

Wechsler, D. (2004). *Wechsler Preschool and Primary Scale of Intelligence* - third edition (*WPPSI-III*). New York: Psychological Corporation.

Westerberg, H., Hirvikoski, T., Forssberg, H., & Klingberg, T. (2004). Visuo-spatial working memory span: a sensitive measure of cognitive deficits in children with ADHD. *Child Neuropsychology*, **10** (3), 155–161.

Wright, S.B., Matlen, B.J., Baym, C.L., Ferrer, E., & Bunge, S.A. (2008). Neural correlates of fluid reasoning in children and adults. *Frontiers in Human Neuroscience*, **1**, 1–8.

Received: 13 March 2010
Accepted: 15 September 2010