Child Neuropsychology, 2011, i*First*, 1–17
http://www.psypress.com/childneuropsych
ISSN: 0929-7049 print / 1744-4136 online
DOI: 10.1080/09297049.2011.575772

Psychology Press
Taylor & Francis Group

Working memory training improves reading processes in typically developing children

Sandra V. Loosli[1,2], Martin Buschkuehl[3], Walter J. Perrig[1], and Susanne M. Jaeggi[3]

[1]Department of Psychology, University of Bern, Bern, Switzerland
[2]Department of Neurology, University Medical Center, University of Freiburg, Freiburg, Germany
[3]Department of Psychology, University of Michigan, Ann Arbor, MI, USA

The goal of this study was to investigate whether a brief cognitive training intervention results in a specific performance increase in the trained task, and whether there are transfer effects to other nontrained measures. A computerized, adaptive working memory intervention was conducted with 9- to 11-year-old typically developing children. The children considerably improved their performance in the trained working memory task. Additionally, compared to a matched control group, the experimental group significantly enhanced their reading performance after training, providing further evidence for shared processes between working memory and reading.

Keywords: Transfer; Scholastic achievement; Cognitive intervention; Literacy; Education.

In everyday life, working memory (WM) is essential for many cognitive processes such as language comprehension, planning or problem solving, and fluid intelligence (e.g., Cowan et al., 2005; Miyake & Shah, 1999). We refer to WM as a dynamic processing system, which is capable of temporarily storing and manipulating a limited amount of information (Kane & Engle, 2002). It is widely assumed that the capacity of WM is fixed to a certain amount of items or chunks, and it has been demonstrated that this amount is age related and varies considerably among individuals (Cowan, 2001; Gathercole, 1999; Pickering, 2001; Rouder et al., 2008). Individual differences in the capacity of WM may arise for several reasons, including limited capacity per se (e.g., Cowan, 2005) or limitations in the cognitive control mechanisms supporting WM (e.g., Hasher, Lustig, & Zacks, 2008; Hasher & Zacks, 1988).

The preparation of this manuscript was partly supported by a grant of the German Federal Ministry of Education and Research (BMBF-01GW0710) to S.V.L. and by Swiss National Science Foundation Fellowships PBBE1-117527 to M.B. and PA001-117473 to S.M.J. We thank Martina Bichsel, Fabia Fischli, Axel Rau, Franziska Schachtler, Anja Scholl, and Nathalie Weber for their help with testing, and Daniela Blaser for her assistance during training sessions. Further, we thank the teachers and children for participating in this study, Kristin Flegal for editorial advice, and Priti Shah and John Jonides for helpful discussions on an earlier draft of this work.

Address correspondence to Susanne Jaeggi, Department of Psychology, The University of Michigan, East Hall, 530 Church Street,Ann Arbor, MI, 48109-1043, USA. E-mail: sjaeggi@umich.edu

It has been argued that WM capacity is crucial for children's general ability to acquire knowledge and new skills (Alloway, Gathercole, Adams, & Willis, 2005; Gathercole, Lamont, & Packiam Alloway, 2006; Gathercole, Pickering, Knight, & Stegman, 2004). There is also evidence that WM capacity is directly related to scholastic achievement (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Pickering, 2006; Rapport, Scanlan, & Denney, 1999), as for example, to mathematical skills (e.g., Alloway et al., 2005; Bull & Scerif, 2001; Gathercole, 1999; Mayringer & Wimmer, 2000; McLean & Hitch, 1999; Passolunghi & Siegel, 2001; Siegal & Ryan, 1989), but also to vocabulary (Daneman & Green, 1986), language comprehension (e.g., Nation, Adams, Bowyer-Crane, & Snowling, 1999; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000), and reading ability (e.g., de Jong, 1998; Gathercole, Brown, & Pickering, 2003; Gathercole et al., 2004; Swanson, 1994).

There is a lot of research specifically dedicated to the investigation of the relationship between reading processes and WM. For example, it has been shown in adults that reading comprehension relies on the executive or attention component of WM but also, although to a lesser extent, on the phonological processing and storage component (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). However, de Jonge and de Jong (1996) demonstrated that both executive processing as well as simple storage are equally related to reading comprehension in typically developing children and that both are related to reading speed. Further evidence for an association between WM and reading comes from studies with dyslexic children, which often have phonological deficits (Goulandris, 2003; Snowling, 1998) and, compared to typically developing children, show poorer performance in complex WM spans (Jeffries & Everatt, 2004; Reiter, Tucha, & Lange, 2004), tasks that tap the executive component of WM (Daneman & Carpenter, 1980).

Looking beyond specific school-related skills, there is also evidence that children with low-WM capacity need additional classroom support in school in order to achieve appropriate goals (Alloway et al., 2009). Furthermore, these children have low attention spans, are easily distracted and tend to forget instructions. Therefore, low-WM capacity seems to be a high-risk factor for underachievement in early school years (Alloway et al., 2009).

In sum, WM seems to be a very important factor for scholastic achievement in many domains, including reading. Consequentially, even small increases in the efficacy of WM may significantly improve children's performance in the classroom and in their daily lives (Minear & Shah, 2006). Therefore, it is not surprising that WM training in children is of high interest. Klingberg and colleagues for example were able to show that WM training with children diagnosed with attention deficit/hyperactivity disorder (ADHD) not only leads to performance improvements in the trained task but also improves other nontrained WM tasks and even matrix reasoning (Klingberg et al., 2005; Klingberg, Forssberg, & Westerberg, 2002). Further, there is some recent evidence by this group that WM training has beneficial effects on nontrained WM tasks and attention in typically developing preschool children (Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2008). Studies with young adults have demonstrated transfer of WM training either within the same cognitive domain (Dahlin, Stigsdotter Neely, Larsson, Backman, & Nyberg, 2008; Persson & Reuter-Lorenz, 2008), or even to more general cognitive abilities relying on WM, such as fluid intelligence (*Gf*; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). What might be the mechanism responsible for those transfer effects? Although the aforementioned studies are quite diverse in terms of samples, training, and transfer tasks, all of them train some aspect of WM, and, furthermore, the transfer is found in domains that are closely related to WM. Thus, it seems that transfer occurs if the training and transfer task share common processes (Jaeggi et al., 2008; Persson & Reuter-Lorenz, 2008). Considering that WM is the underlying process that determines the performance of a multitude of tasks (Cowan et al., 2005), one could predict that training WM will be beneficial for all other tasks and domains that rely on the efficacy of WM. Direct evidence for this view comes from a recent study demonstrating that transfer occurs if the training and the transfer task engage overlapping brain regions, but not if they engage different regions (Dahlin et al., 2008).

Despite the importance of WM for scholastic achievement, there are very few studies to date that examined effects of WM training on standardized academic achievement tests. Traditionally, interventions aimed at reading difficulties by primarily focusing on phonological awareness (e.g., Brady, Fowler, Stone, & Winbury, 1994) or other phonological processes (e.g., Cohen et al., 2006) rather than on WM processes. Early, noncomputerized training studies with children that focused directly on WM provided mixed results but indicated that there is some potential to improve reading by training WM (Gutezeit & Meier, 1977; Hays & Pereira, 1972; Maridaki-Kassotaki, 2002; Whisler, 1974). Recently, St. Clair-Thompson, Stevens, Hunt, and Bolder (2010) examined a computer game that was designed to train memory strategies in typically developing children. Although the intervention group improved in various WM tasks, no gains in standardized academic tests such as reading or arithmetic were obtained. However, there is considerable evidence that strategy training yields large improvements in the trained task but usually results in only small transfer effects due to the intense practice of very task-specific skills (Lustig, Shah, Seidler, & Reuter-Lorenz, 2009). Holmes, Gathercole, and Dunning (2009) trained low-WM children with adaptive WM tasks as used by Klingberg et al. (2005) and observed improvements in several nontrained WM tasks. However, also in this study, there was no evidence for training-related improvements in reading, mathematical ability, or in general intelligence. In contrast, another recent study showed larger improvements in reading speed after a computerized WM training in adult dyslexic readers than in age-matched nondyslexic readers (Horowitz-Kraus & Breznitz, 2009). Finally, Chein and Morrison (2010) trained college students on a visual and a verbal complex span task for 4 weeks and observed transfer on a reading comprehension task.

To conclude, evidence for improved reading after WM training is very scarce, and there are no studies available investigating whether WM training improves reading processes in typically developing children. Therefore, the goal of the current study is to determine whether WM training in this group will lead to transfer effects to important school-related domains; in our case, reading performance.

For the present study, we developed a special variant of a WM task that we used previously in order to train old adults (Buschkuehl et al., 2008). This training taps both storage and processing resources and is based on classic WM span measures such as the reading, operation, or spatial span tasks (Conway et al., 2005; Daneman & Carpenter, 1980; Shah & Miyake, 1996). In our version, participants had to identify the orientation of each picture in a series of animal pictures; that is, they had to decide whether an animal was presented correctly or upside-down. Immediately after presentation of the series, the children had to recall the order in which the animals were presented. Since verbal rehearsal strategies are already common at the age of our sample (Gathercole, 1999), the response window after the presentation of each item was kept short in order to minimize the use of verbal rehearsal or other strategies. Moreover, the training task was adaptive, forcing the participants to train at their WM capacity limit: After each successful trial, WM load was increased by one item, and after failure, it was reduced by one item.

We used three different measures for reading ability as obtained with the Salzburg Reading Test (SLT; Landerl, Wimmer, & Moser, 1997) in order to investigate differential effects of training: Pseudowords, single words, as well as text passages. Synthetic and phonological reading are assessed with pseudoword reading, requiring self-generated pronunciation without the possibility to access previously acquired and stored knowledge. Automatic and direct word recognition is assessed with single and composite words. The third part consists of reading a short, coherent story (Landerl et al., 1997). The requirement is to read as quickly and accurately as possible in all subcomponents. As we train on WM, we expected more transfer to those reading tasks that place more demands on WM. Thus, we predicted the largest transfer effect for text reading because proficient reading requires the simultaneous engagement of multiple WM-related processes: Information must be held in WM until a sentence has been read to the end. Additionally, reading demands the continuous encoding of words, the storage of these words and/or the general content across sentences, building inferences between words and sentences, and, finally, comprehension and interpretation (Just & Carpenter, 1992). These multiple processes are less important in single word and pseudoword reading and, thus, our WM training should have a smaller impact on those tasks.

In addition to reading performance, we also included a transfer task, which is highly correlated with measures of scholastic achievement, namely *Gf* as measured with a matrix reasoning task. Concerning *Gf*, the dose-response curve observed in Jaeggi et al. (2008) led to the prediction of only small transfer effects (if any), since the training time in the present study was considerably shorter than the one implemented in the previous study. In addition, the training task we used here does not involve as much visuospatial WM processes as the one used in Jaeggi et al. (2008). Consequently, there are less overlapping processes between the training and the transfer task. Thus, although we did not expect large effects in *Gf* in the present study, the inclusion of this task is important in order to get a sense of how this training compares to our successful dual *n*-back training.

Furthermore, differential effects of training and transfer were of interest. One important issue is to ensure that any observed transfer effect is strictly training related and not due to preexisting individual differences. Therefore, appropriate matching between the training and the control group is essential. Nevertheless, within the training group, we have seen that individual differences in intelligence affect the transfer results in that participants with initially lower intelligence scores showed larger transfer effects (Jaeggi et al., 2008), presumably because there was more room for improvement. Thus, also in this study, we tested whether interindividual differences as assessed at pretest are related to training performance and transfer. Additionally, we examined whether and how performance in the training task affects transfer.

The present study differs in various ways from previous training research conducted with children. Unlike many other WM-training studies in which different tasks were combined into multidomain interventions (e.g., Buschkuehl et al., 2008; Klingberg et al., 2002, 2005; Persson & Reuter-Lorenz, 2008; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005), we used only one training task, making sure that the effects could be traced back to one specific task and its involved processes. Also, we allocated only a minimal amount of training time (10 sessions), which is considerably shorter than the training time used in other approaches (e.g., Klingberg et al., 2002, 2005). Finally, we used an adequate number of subjects and an adequate, well-matched control group.

METHOD

Participants

A total of 66 children aged 9 to 11 ($M = 9.50$ years, $SD = 0.56$) participated in the study. All of them attended the end of third grade or the beginning of fourth grade in elementary schools near Bern, Switzerland, at the time of data collection. The parents gave written consent for their children to participate in the study. The experimental group consisted of one entire class of 24 children, and 42 children from four other classes formed the control group. No payment or rewards were offered for participation.

After data collection, 6 children (4 from the experimental group) were excluded from data analyses because of known and diagnosed neurological, psychiatric, or developmental disorders as assessed by teacher reports. Three of them had diagnosed developmental disorders (one ADHD, one visual perception disorder, and one auditory perception disorder), and for 1 child there was reasonable evidence for an epileptic disorder. Another child had a concussion between pre- and posttest sessions, and, finally, 1 child was excluded because of missing data in the pretest session due to time constraints. The remaining 40 children of the control group were matched one by one with a child of the remaining 20 children of the experimental group; thus the data of 20 control children were not used for data analyses. Matching criteria were gender, age in months, and pretest scores in *Gf*. Additional factors such as special education needs or native language were held constant between groups as much as possible. Most of the children were native Swiss-German speakers. Of the non-native Swiss-German-speaking children, all of them understood and spoke German fluently without having any problems following the training and test sessions. After this matching procedure, both groups consisted of 20 children with 11 girls and 9 boys in each group. The mean age was 119.75 months ($SD = 5.95$) for the experimental group, and 120.25 months ($SD = 4.95$) for the control group, $t(38) = 0.29$; $p = ns$. Concerning *Gf* at pretest, there was no significant group difference either (experimental group: $M = 24.16$, $SD = 4.32$; control group: $M = 26.30$, $SD = 4.88$), $t(38) = 1.47$; $p = ns$.

Apparatus

The training was performed using Windows-based computers with 17-inch screens set at a 1024 × 768 resolution. The training was programmed with E-Prime (Release 1.1) and was conducted in a computer lab consisting of 12 machines. No previous computer skills were required for the completion of the training. All transfer tasks were administered as paper-and-pencil versions.

Task Design and General Procedure

The experimental group received 2 weeks of WM training, preceded and followed by pre- and posttests. The control group participated in the pre- and posttests at the same test-retest interval but did not take part in any intervention.

The pre- and posttests were conducted 3 or 4 days before and after the training, respectively. Thus, in the experimental group, the time between pre- and posttests was 18 days, except for 2 children for whom the time was 19 days. In the control group, the time between both test sessions was 18 days for all children.

The training took place over 2 consecutive weeks from Monday to Friday during the first school lesson in the morning. For the training, the class was divided into two groups of 12 children who performed the training individually in a computer lab. The group to go first alternated each day. The training consisted of two 6-minute blocks separated by a 5-minute break.

Material and Procedure

Training The training task was a WM span task consisting of two parts, as illustrated in Figure 1. In the first part (processing/encoding stage), a sequence of animals was presented in the center of the computer screen. There were four possible animals (300 × 300 pixels in size), which could be named with two syllables in Swiss-German (camel,

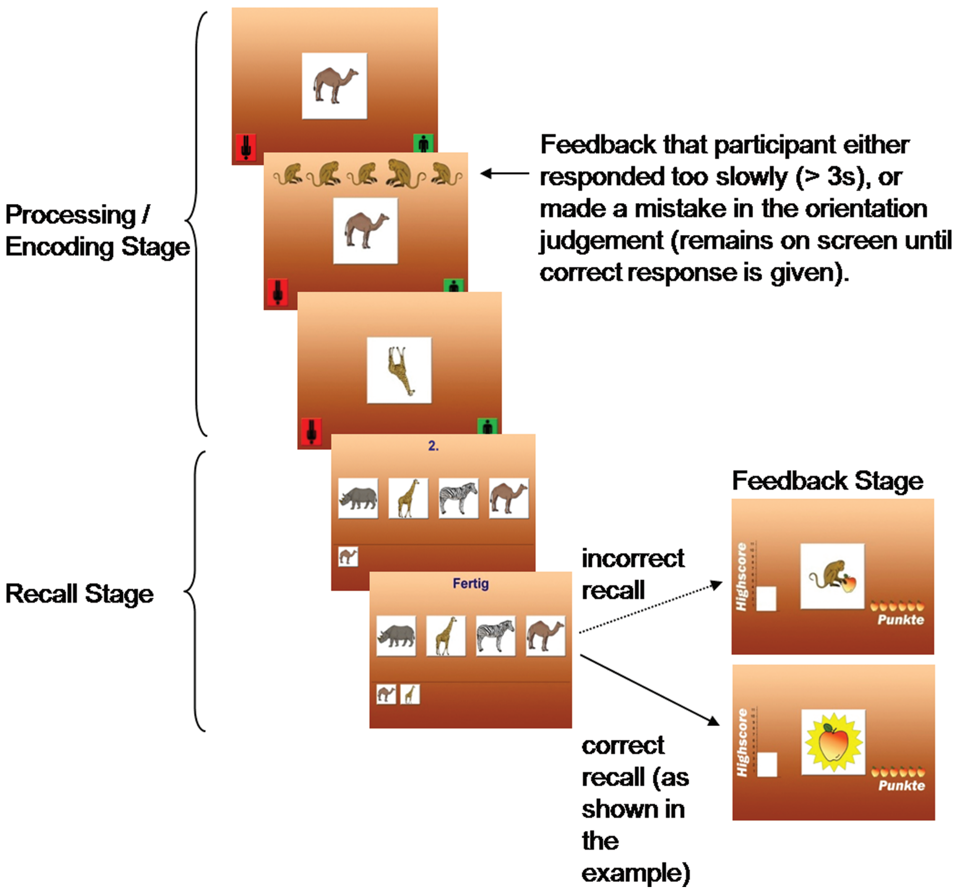

Figure 1 Training task: A complete trial with a set size of two is shown. In the encoding/processing stage, participants have to decide about the orientation of the picture (click on the green mouse button [right] for correct presentation or on the red button [left] for upside-down) while encoding the sequence at the same time. Five monkeys appear if participants make a mistake in this stage or if they take more than 3 s to respond. In the recall stage, participants have to reproduce the sequence of animals presented in the encoding/processing stage by clicking on the appropriate animals from the display in the correct order. At the end of the trial, participants receive feedback on whether the recall was correct or not. In addition, the current high score is given on the left, indicating the highest set size obtained in a particular session so far. After the feedback screen, a new trial begins with the set size depending on performance in the previous trial (see text for details).

zebra, giraffe, and rhino). The stimuli were taken from the Snodgrass and Vanderwart-"like" objects (Rossion & Pourtois, 2004) consisting of colored and shaded line drawings. Each picture appeared either normally or upside-down, and the participants had to decide on the orientation of the picture by pressing the appropriate mouse button (right/green for correct presentation, left/red for upside-down). At the same time, the sequence of animals had to be encoded in order.

Participants were allowed 3000 ms to give their orientation answer; otherwise, five monkeys appeared on the upper part of the computer screen (see Figure 1) as a prompt to respond quickly. The same monkeys also appeared if the participants hit the wrong button.

In the second part (recall stage), participants had to reproduce the previously shown animal sequence in the correct order. They could do this by clicking on the appropriate pictures (200×200 pixels) consisting of the four possible animals that were displayed in the upper half of the screen (see Figure 1; recall stage). As a confirmation for their mouse click, a small picture (100×100 pixels) of the chosen animal appeared at the bottom of the screen. Once an animal was chosen by a mouse click, it could not be changed, but there was no time limit for the recall of the items.

The orientation of the animals in the first part of the task was randomly assigned. The sequence of animals was also determined randomly and, thus, each animal could appear multiple times within a sequence; however, there was the restriction that no animal could be presented more than twice in a row.

An important feature of the training was its adaptivity in that task difficulty was matched to the actual performance of each individual: If the subjects made no mistakes in the first part and reproduced the previously shown sequence correctly in the second part, the next sequence length (henceforth termed "set size") was increased by one animal. If the participants made a mistake in the first part (i.e., by pressing the wrong mouse button, taking too much time to answer, or both) but reproduced the sequence correctly, the next set size remained unchanged. If the children reproduced the sequence incorrectly (regardless of the performance in the first part) the next set size was reduced by one animal. The smallest possible set size was two.

In order to make the training task attractive for children, a performance feedback was given at the end of each trial (see Figure 1): Participants could win points that were expressed as apples that piled up. Each time the presented sequence of animals was reproduced in the correct order, the participant won an apple. If the repeated sequence was incorrect, no apples were earned, and, in addition, a monkey appeared that prevented the participant from winning an apple. However, apples that were already won were not taken away. Furthermore, a performance indicator was given as a high score representing the maximum number of animals that could be reproduced in the correct order if there was no mistake in the first part of the training task. After the feedback screen, the next trial started as soon as the participant hit the mouse button.

In each training session, the task started with a set size of two animals and ended after 6 minutes. The averaged set size per run served as the dependent variable.

Transfer tasks

Test of Nonverbal Intelligence (TONI; Brown, Sherbenou, & Johnsen, 1997) This task was used to measure abstract, figural problem solving, that is fluid intelligence (*Gf*; Brown et al., 1997). Out of four to six alternatives, the child had to choose the solution that logically concluded the pictures of geometric patterns depicted above the response alternatives. The participants were given a test booklet with 45 problems. Due to time constraints, the time for task completion was restricted to 10 minutes in which the children had to solve as many problems as possible. The children were told that they might not reach the end of the booklet within this time, but that they should just try to solve as many problems as possible. The standardized procedure according to the test manual gives no time limit to complete the test (Brown et al., 1997). However, most of the children completed all problems in 10 minutes or less, thus we think that this procedure had no adverse effects on the validity or reliability of the test as a measure of *Gf*. Further, we only used raw scores instead of the norm-based data for our analyses. Before the actual task, five practice trials were given in order to ensure that the task was properly understood. The dependent variable consisted of the number of correctly solved trials in this time frame. Parallel forms were used for the pre- and posttests.

Salzburger Lesetest (SLT) This task is the reading part of the Salzburg Reading and Spelling Test and is widely used to assess reading ability in German-speaking countries (Landerl et al., 1997). The correlations between the subtests used here and German grades range between $r = .37$ and $r = .54$ (Landerl et al., 1997), thus, it seems that the test has some external validity, and we consider it as a valid predictor of school-relevant skills. The parallel test reliability scores for reading speed in this test are adequate (between $r = .84$ and $r = .99$; Landerl et al., 1997) and therefore suitable for pre- and posttests in a training study.

The children were given five different sheets containing pronounceable pseudowords (24 dissimilar to words, 30 similar to words), words (30 frequent words, 11 compound words), as well as a short and meaningful story (57 words long) in which context facilitated successful reading, representing a normal reading situation (Landerl et al., 1997). Words, pseudowords, and story had to be read as quickly and as correctly as possible and reading times and errors for the whole sheet were recorded. As dependent variables, the RTs and number of errors were z-transformed and summarized into three composite measures, one for pseudowords, one for words, and one for text. Parallel forms were used in the pre- and posttest sessions.

Data Analysis

Statistical analyses were performed with SPSS (Release 15.0.1). All statistic tests are based on a significance level of $\alpha = .05$.

For the experimental group, the mean training level was assessed for each child and training session, and the mean set size per training day was calculated. In order to get a measure for the individual training gain, we calculated the individual performance differences between the first and the last training day. Additionally, the mean set size over all sessions was calculated. For all training data, the first three trials of each session were regarded as warm-up trials and were excluded from data analyses. In cases of missing data, the respective value was interpolated between the former and latter session based on the individual training scores. However, this affected only two values, which represented less than 0.01% of all data points.

We also investigated whether interindividual differences were related to training performance by means of correlational analyses. We considered performance on the intelligence and reading measure at pretest as potential determinants of training and transfer.

Thus, we correlated these variables with the overall training performance (mean all training sessions), and the training gain (difference of the last and the first training session, divided by the first training session).

For the transfer measures, we first tested whether the two training groups differed at pretest using independent t-tests. We analyzed a global transfer effect by calculating a multivariate analysis of variance (MANOVA) with group (experimental, control) as the between factor and the differences between post- and pretest scores (from this point on termed gain scores) as dependent variables. Following the suggestions by Olson and Stevens (Olson, 1974, 1979; Stevens, 1979) we report Pillai's V as an F-statistic, assuming that it yields the most robust outcome. Further, we used a discriminant function analysis and separate independent t-tests with the gain scores as dependent measures to follow up the MANOVA.

Finally, we tested whether interindividual differences were related to the transfer effects. We took the variables in which we found significant transfer effects by means of the MANOVA and t-tests as described above and correlated them with *Gf* and reading performance at pretest, as well as the overall training performance and the training gain.

RESULTS

Specific Training Effect

The mean performance across all children is depicted in Figure 2. There were large interindividual differences in training performance, as there were children who highly improved their performance, while others showed almost no improvement at all. On average, the children performed 55.6 trials ($SD = 9.13$) per training day and reached an overall set size of 4.03 ($SD = 0.61$). The highest set size consisted of eight items, which was only reached by four children. On the first day, the children started on an average set size of 3.37 ($SD = 0.57$). On the last day, their average set size was 4.16 ($SD = 0.84$), representing a significant performance increase of 23%, $t(19) = -6.20$, $p < .001$, $d = 1.09$. The increase from the second ($M = 3.92$, $SD = 0.63$) to the last training day also remained significant (6% increase), $t(19) = -1.97$, $p < .05$, $d = 0.32$; however, comparing the third ($M = 4.02$,

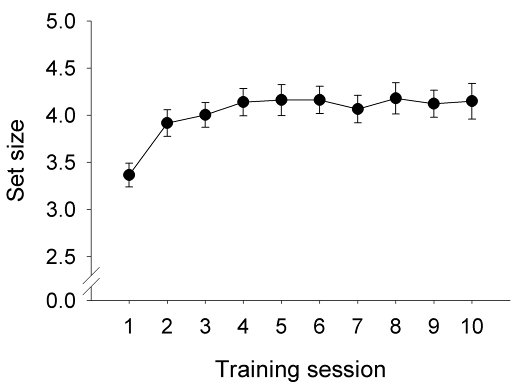

Figure 2 Training performance for all participants. The mean training levels (average set size) as well as the corresponding standard errors of the mean are depicted for each training session. The first three trials of each training session were excluded from the analysis.

Table 1 Pearson's Correlations Between Interindividual Differences and Training Performance.

	TONI pretest	SLT pretest	Overall training
Fluid intelligence (TONI) at pretest			
Reading (SLT) at pretest	.06		
Overall training performance (mean all sessions)	.21	−.33	
Training gain (session 10 minus Session 1)	−.02	−.38	.60**

Note. $n = 20$. TONI = Test of Nonverbal Intelligence. SLT = Salzburger Lesetest (composite score all subtests).

$^{**}p < .01$.

$SD = 0.59$) with the last day, the improvement was no longer significant (3% increase), $t(19) = -1.33$, $p = .10$, $d = 0.19$.

Interindividual Differences and Training Performance

The correlation matrix correlating training performance (overall training performance and training gain) and the interindividual differences variables (*Gf* and reading at pretest) is displayed in Table 1.

Transfer Effects

Descriptive data for pre- and posttest scores for each group, as well as the effect sizes for the pre-post differences, are reported in Table 2. There were no significant differences between groups in any transfer measures at pretest.

Table 2 Descriptive Data (Mean and Standard Deviations) in Pre- and Posttest Sessions for the Experimental and Control Group, as Well as Difference Scores and Effect Sizes (Cohen's *d*) for the Pre- and Posttest Comparisons for Each Group.

			Pretest		Posttest			
			M	*SD*	*M*	*SD*	*Difference Score*	*d*
Fluid Intelligence (TONI)								
		EG	24.16	−4.32	25.45	−3.76	1.29	0.32
		CG	26.30	−4.88	27.05	−3.17	0.75	0.18
Reading Test (mean RT in seconds & errors)								
Pseudowords	Time	EG	39.43	10.95	39.62	14.27	0.19	−0.01
		CG	41.81	11.69	41.33	10.73	−0.48	0.04
	Errors	EG	2.00	1.18	1.78	1.38	−0.22	0.17
		CG	1.90	2.07	2.38	2.80	0.48	−0.19
Words	Time	EG	26.30	7.13	21.09	7.17	−5.21	0.73
		CG	24.26	8.86	21.37	9.81	−2.89	0.31
	Errors	EG	0.90	0.72	0.48	0.57	−0.42	0.65
		CG	0.60	0.94	0.35	0.63	−0.25	0.31
Text	Time	EG	38.52	9.34	36.57	10.32	−1.95	0.20
		CG	34.64	9.03	36.59	10.72	1.95	−0.20
	Errors	EG	1.65	2.18	1.15	1.42	−0.50	0.27
		CG	0.60	1.14	1.45	2.21	0.85	−0.48

Note. EG: Experimental group; CG: Control group; TONI: Test of Nonverbal Intelligence. Experimental group: $n = 20$, Control group: $n = 20$.

Table 3 Standardized Canonical Discriminant Function Coefficients and Canonical Variate Correlation Coefficients for the Significant Function of the Transfer Measures.

Dependent variable	Standardized canonical discriminant function coefficients	Canonical variate correlation coefficients
Fluid Intelligence (TONI)	.26	−.11
Reading gain – Pseudowords	.17	.14
Reading gain – Words	.55	.56
Reading gain – Text	.87	.80

Note. TONI = Test of nonverbal intelligence.

The MANOVA with all outcome measures (*Gf*, reading of pseudowords, words, and text) as dependent variables was significant, $F(4, 35) = 3.80$, $p < .05$, $\eta_p^2 = .30$. The discriminant function analysis resulted in one function that was significant, Wilks' Lambda = .70, χ^2 (4, $N = 40$) = 12.99, $p < .05$, and comprised all four outcome measures. Table 3, with the standardized canonical discriminant function coefficients and the canonical variate correlation coefficients, indicates that the function is mainly driven by the gain in the text-reading measure and to a lesser degree by the word-reading measure. Smaller contributions resulted from the *Gf* measure and the pseudoword-reading. The findings from the independent *t*-tests showed that the trained group significantly improved more than the control group in text reading, $t(38) = -3.25$, $p < .01$, $d = 1.03$, as well as in word reading, $t(27.87) = -2.28$, $p < .05$, $d = 0.72$, but not in the other measures ($d < 0.18$).

Interindividual Differences and Transfer

None of the significant transfer effects (gain in word and text reading) were significantly correlated with the variables we considered here (pretest performance in *Gf* [TONI] or reading [SLT], overall training performance, or training gain).

DISCUSSION

In this study, 20 children were trained with a WM-based training intervention for two weeks. Children were tested on measures of reading and problem solving before and after the training. Their performance was compared to a matched control group that did not engage in any training activity. Our participants not only showed a significant gain in the trained task but also in untrained measures, in particular, reading of single words and text.

The specific training data showed that the children significantly improved their performance in the trained task. The largest improvement occurred in the first four training days, after which the performance increase leveled off. This training curve is different from some previous studies, in which participants showed a linear performance increase (Chein & Morrison, 2010; Dahlin et al., 2008; Jaeggi et al., 2008; Klingberg et al., 2005). But, on the other hand, there are also studies in which participants show almost no training gain throughout the intervention (Buschkuehl et al., 2008, Figure 3c; Olesen, Westerberg, & Klingberg, 2004, Exp. 2; Thorell et al., 2008). Nevertheless, despite the different training curves, all of the aforementioned studies reported transfer effects to untrained tasks. Therefore, it would be of interest for future studies to investigate whether and how the quality of the training is related to the amount of transfer.

Concerning the transfer measures, we found an overall larger performance increase in the experimental group as indicated by the MANOVA. The MANOVA was driven by the gain in reading performance, that is, in reading of text and words but not pseudowords. What might be the underlying mechanisms for such transfer? In terms of the processes involved, we think that our transfer task shares important features with the training task that we defined as an essential prerequisite to obtain transfer effects (Jaeggi et al., 2008). The training task is a classic complex span task. Such tasks often have been shown to be related to reading comprehension (Daneman & Merikle, 1996) but also to reading speed (de Jonge & de Jong, 1996). The differential transfer effects in the reading subtests suggest that the training only affected those reading processes that demand a considerable amount of WM and that require access to lexical knowledge. This is consistent with prior research showing that there is a stronger link between verbal WM and familiar lexical items than for unfamiliar nonword stimuli (Gathercole, Pickering, Hall, & Peaker, 2001; Roodenrys, Hulme, & Brown, 1993). This seems to be in contrast to neuroimaging data showing greater activity in inferior frontal lobe structures during reading or rehearsal of pseudowords relative to familiar words (e.g., Fiez et al., 1996; Mechelli, Gorno-Tempini, & Price, 2003). However, pseudowords may increase the demands for lexical processing because there is no semantic representation available (Mechelli et al., 2003) and could hence be related to effort (Fiez, Balota, Raichle, & Petersen, 1999). Thus, this activation pattern does not necessarily provide evidence for a stronger association between WM and pseudowords than between WM and familiar words. Furthermore, there is evidence for a direct relationship between WM performance and vocabulary knowledge in children (e.g., Gathercole & Baddeley, 1993; Jarrold, Baddeley, Hewes, Leeke, & Phillips, 2004). Our results are also consistent with findings by Unsworth and Engle (2006) showing that errors in a complex span task are related to memory retrieval. It is possible that our training has improved retrieval processes, which in turn enhanced reading speed of single words and text with semantic content, since decoding performance relies on retrieval of vocabulary (Perfetti, 2010). As we found the strongest transfer effects in text reading, one could argue that in text reading as opposed to single-word reading, additional syntactic processes come into play, which strongly rely on working memory capacity (cf. Caplan & Waters, 1999). This is in line with research that shows a stronger relationship between WM and the comprehension of complex sentences than WM and simple sentence comprehension (Montgomery, Magimairaj, & O´Malley, 2008). Alternatively, it has been argued that the strong relationship between WM and reading primarily results from the involvement of attention control in both tasks (e.g., Hasher et al., 2008; Hasher & Zacks, 1988). Thus, in more general terms, our training procedure may have facilitated the ability to control attention.

Interestingly, although there was considerable variance in terms of training performance, we could not pinpoint a specific feature from training performance that correlated with transfer to reading, and, further, neither performance on the intelligence measure nor reading performance at pretest correlated with the transfer measures. Although a somewhat surprising finding, it could well be that there are other training-related features that are related to transfer but that we did not capture, such as, for example, attentional or motivational factors. In addition, we think that, due to the relatively small sample size, the power to detect such differences might have been limited.

Due to the short training time, we did not expect large effects on *Gf* (cf. Jaeggi et al., 2008), also since two other studies that trained ADHD children observed transfer effects on *Gf* only after 5 weeks involving sessions of 40 minutes each (Klingberg et al., 2002, 2005).

In addition, the same group failed to show transfer on *Gf* with a shorter training (Thorell et al., 2008). Thus, considering that our training intervention was merely 10 sessions long, our lack of transfer to *Gf* is hardly surprising; although there is now recent evidence that transfer to *Gf* is possible with very little training time (Karbach & Kray, 2009). Our results, however, are comparable to those of Chein and Morrison (2010), who also trained their participants on a complex WM task and found no transfer to *Gf.*

Although our data are well in line with widely accepted theoretical assumptions regarding the relationship between WM and reading performance, our study has some limitations that have to be considered: First of all, we used a quasi-experimental design that comes with its associated disadvantage that there might be some inherent group differences. Due to our matching procedure, we observed no significant differences in the pretest scores, but nonetheless there could have been other differences in the participating classes that we did not capture. Related to this issue, the fact that we did not include an active control group to control unspecific effects of the training might suggest that the observed transfer effects are driven by some unspecific factors (e.g., Hawthorne effect). However, as we have found reliable transfer effects using the same task in other participant groups before (Buschkuehl et al., 2008), we rather think that our effects are driven by the training intervention. If the gain of the experimental group was mainly based on such unspecific factors, we should have seen advantages for the experimental group in all pre- and posttest measures, which was not the case.

Finally, we are aware that the SLT does not measure the full range of reading-related processes as it does not assess comprehension, for example, but, nevertheless, we think that this test is a good proxy for measuring basic reading processes, and, also, it seems to be a valid predictor of school-relevant skills (Landerl et al., 1997). These caveats notwithstanding, our results are quite remarkable given the brevity of the intervention and given its impact on cognitive processes, which are of importance for scholastic achievement. Additionally, our results provide further evidence that WM and reading share overlapping processes. Finally, in contrast to many previous WM training studies, which often looked at transfer on other laboratory tasks, we showed that it is possible to improve an ability that is very important in everyday life and is related to scholastic achievement in school-aged children.

Original manuscript received September 6, 2010
Revised manuscript accepted March 5, 2011
First published online

REFERENCES

Alloway, T. P., Gathercole, S. E., Adams, A.-M., & Willis, C. (2005). Working memory abilities in children with special education needs. *Educational and Child Psychology, 22*(4), 56–67.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*(2), 606–621.

Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia, 44*, 26–59.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *TONI-3: Test of Nonverbal Intelligence* (3rd ed.). Austin, TX: Pro-Ed.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*(3), 273–293.

Buschkuehl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Dapp, C., Muller, M., et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology and Aging, 23*(4), 743–753.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77–94; discussion 95–126.

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17*(2), 193–199.

Cohen, D., Plaza, M., Perez-Diaz, F., Lanthier, O., Chauvin, D., & Hambourg, N. (2006). Individual cognitive training of reading disability improves word identification and sentence comprehension in adults with mild mental retardation. *Research in Developmental Disabilities, 27*, 501–551.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–185.

Cowan, N. (2005). *Working memory capacity*. New York, NY: Psychology Press.

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*(1), 42–00.

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

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25*, 1–18.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433.

de Jong, P. F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology, 70*(2), 75–96.

de Jonge, P., & de Jong, P. F. (1996). Working memory, intelligence and reading ability in children. *Personality and Individual Differences, 21*(6), 1007–1020.

Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron, 24*, 205–218.

Fiez, J. A., Raife, E. A., Balota, D. A., Schwarz, J. P., Raichle, M. E., & Petersen, S. E. A. (1996). Positron emission tomography study of the short-term maintenance of verbal information. *Journal of Neuroscience, 16*, 808–822.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410–419.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education, 8*, 259–272.

Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of national curriculum attainment levels. *Educational and Child Psychology, 20*, 109–122.

Gathercole, S. E., Lamont, E., & Packiam Alloway, T. (2006). Working memory in the classroom. In S. Pickering (Ed.), *Working memory and education* (pp. 219–240). Oxford, UK: Elsevier Press.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology A, 54*(1), 1–30.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegman, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1–16.

Goulandris, N. (2003). Introduction: Developmental dyslexia, language and orthographies. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 1–14). London, UK: Whurr.

Gutezeit, G., & Meier, E. (1977). Zur Effektivität eines projektionstachistoskopischen Übungsprogramms bei legasthenen Kindern aus 4. Grundschulklassen. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 26*(7), 266–274.

Hasher, L., Lustig, C., & Zacks, R. (2008). Inhibitory mechanisms and the control of attention. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 227–249). Oxford, NY: Oxford University Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York, NY: Academic Press.

Hays, B. M., & Pereira, E. R. (1972). Effect of visual memory training on reading ability of kindergarten and first grade children. *Journal of Experimental Education, 41*(1), 33–38.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12*(4), F9–F15.

Horowitz-Kraus, T., & Breznitz, Z. (2009). Can the error detection mechanism benefit from training the working memory? A comparison between dyslexics and controls — An ERP study. *PLoS ONE, 4*, 1–10.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105*(19), 6829–6833.

Jarrold, C., Baddeley, A. D., Hewes, A. K., Leeke, T. C., & Phillips, C. E. (2004). What links verbal short-term memory performance and vocabulary level? Evidence of changing relationships among individuals with learning disability. *Journal of Memory and Language, 50*, 134–148.

Jeffries, S., & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia, 10*, 196–214.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review, 99*, 122–149.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin & Review, 9*(4), 637–671.

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.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (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.

Landerl, K., Wimmer, H., & Moser, E. (1997). *Salzburger Lese- und Rechtschreibtest (SLRT). Verfahren zur Differentialdiagnose von Störungen des Lesens und Schreibens für die 1. bis 4. Schulstufe*. Bern, Switzerland: Hans Huber.

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review, 19*(4), 504–522.

Maridaki-Kassotaki, K. (2002). The relation between phonological memory skills and reading ability in Greek-speaking children: Can training of phonological memory contribute to reading development? *European Journal of Psychology of Education, 17*(1), 63–73.

Mayringer, H., & Wimmer, H. (2000). Pseudoname learning by German-speaking children with dyslexia: Evidence for a phonological learning deficit. *Journal of Experimental Child Psychology, 75*, 116–133.

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

Mechelli, A., Gorno-Tempini, M. L., & Price, C. (2003). Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience, 15*(2), 260–271.

Minear, M., & Shah, P. (2006). Sources of working memory deficits in children and possibilities for remediation. In S. Pickering (Ed.), *Working memory and education* (pp. 274–307). Oxford, UK: Elsevier Press.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory. Mechanisms of active maintenance and executive control*. Cambridge, UK: Cambridge University Press.

Montgomery, J. W., Magimairaj, B. M., & O´Malley, M. H. (2008). Role of working memory in typically developing children's complex sentence comprehension. *Journal of Psycholinguistic Research, 37*, 331–354.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139–158.

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

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association, 69*(348), 893–908.

Olson, C. L. (1979). Practical considerations in choosing a MANOVA test statistic: A rejoinder to Stevens. *Psychological Bulletin, 86*(6), 1350–1352.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80*, 44–57.

Perfetti, C. (2010). Decoding, vocabulary, and comprehension. The golden triangle of reading skill. In M. G. McKeown & L. Kucan (Eds.), *Bringing reading research to life* (pp. 291–302). New York, NY: Guilford.

Persson, J., & Reuter-Lorenz, P. A. (2008). Gaining control: Training executive function and far transfer of the ability to resolve interference. *Psychological Science, 19*(9), 881–889.

Pickering, S. J. (2001). The development of visuo-spatial working memory. *Memory, 9*, 423–432.

Pickering, S. (Ed.). (2006). *Working memory and education*. Oxford, UK: Elsevier Press.

Rapport, M. D., Scanlan, S. W., & Denney, C. B. (1999). Attention-deficit/hyperactivity disorder and scholastic achievement: A model of dual developmental pathways. *Journal of Child Psychology and Psychiatry, 40*(8), 1169–1183.

Reiter, A., Tucha, O., & Lange, K. W. (2004). Executive functions in children with dyslexia. *Dyslexia, 11*, 116–131.

Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology, 56*(3), 431–442.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33*, 217–230.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Science of the United States of America, 105*(16), 5975–5979.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102*(41), 14931–14936.

Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing, 13*, 81–103.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*(1), 4–27.

Siegal, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973–980.

Snowling, M. J. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology and Psychiatric Review, 3*(1), 4–11.

St. Clair-Thompson, H., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology, 30*(2), 203–219.

Stevens, J. (1979). Comment on Olson: Choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin, 86*(2), 355–360.

Swanson, H. L. (1994). Short-term memory and working memory—Do both contribute to our understanding of academic achievement in children and adults with learning disabilities? *Journal of Learning Disabilities, 27*, 34–50.

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2008). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*(1), 106–133.

Unsworth, N., & Engle, R. W. (2006). A temporal-contextual retrieval account of complex-span: An analysis of errors. *Journal of Memory and Language, 54*, 346–362.

Whisler, N. G. (1974). Visual-memory training in first grade: Effects on visual discrimination and reading ability. *Elementary School Journal, 75*(1), 50–54.