Can We Enhance Domain-General Learning Abilities to Improve Language Function?

Christopher M. Conway, Michelle A. Gremp, Anne D. Walk, Althea Bauernschmidt and David B. Pisoni

One of the most exciting recent developments in the cognitive neurosciences is the growing realization that the mind and brain are highly plastic. Whereas traditionally the brain was thought to be relatively immutable past a certain age, there is a growing body of evidence suggesting that neural connections remain modifiable even late into adulthood (Kleim & Jones, 2008; van Praag, Kempermann, & Gage, 2000). In addition, recent research has demonstrated that even relatively short-term training can lead to improvements to neurocognitive abilities such as working memory capacity (Klingberg, 2010), with improvements transferring to a host of non-trained tasks of memory and cognition.

In addition to working memory (WM), it may be fruitful to explore whether it is possible to enhance statistical learning abilities. Here we use the term statistical learning in a fairly broad sense, to refer to incidental learning that results in sensitivity to structured patterns in the environment. Under this definition, statistical learning is related to other forms of non-declarative pattern learning abilities such as implicit learning (Perruchet & Pacton, 2006) and sequential learning (Conway, in press). These kinds of domain-general learning abilities appear to be important for language acquisition and processing (Conway & Pisoni, 2008; Gervain & Mehler, 2010; Gogate & Hollich, 2010; Gupta & Dell, 1999; Kuhl, 2004; Reber, 1967; Saffran, 2003; Ullman, 2004). The question we address is whether it is possible to improve statistical learning abilities by capitalizing on the highly plastic nature of the mind and brain, in a similar vein to the demonstrations of improvements to WM capacity. Given that statistical learning is important for language acquisition and processing, then we should expect that if statistical learning can be enhanced through some type of training regimen, that this would result in a better facility for acquiring and processing language.

Providing a demonstration of the causal effects of enhanced statistical learning on language acquisition would be important theoretically as well as clinically. A number of language and communication disorders in fact may be due at least in part to disturbances to domain-general learning abilities such as statistical learning and procedural memory (Nicolson & Fawcett, 2007; Ullman & Pierpont, 2005) including dyslexia (Howard, Howard, Japikse, & Eden, 2006), specific language impairment (Evans, Saffran, & Robe-Torres, 2009), and language delays caused by a period of deafness early in development (Conway, Pisoni, Anaya, Karpicke, & Henning, 2011).

In this chapter, we first review recent evidence highlighting the importance of statistical learning for language in populations both with and without a language or communication disorder. We then describe recent work using computerized training techniques that were designed to improve WM. This provides the background for presenting the results of a novel adaptive training task that we have developed to improve domain-general learning abilities. We review two studies still in their formative stages, the first with normal-hearing adults, the second with children who are deaf or hard of hearing. The initial findings, although still preliminary, show the promise of adaptively training basic elementary mechanisms of learning and memory to improve language and communication functions.

1. Statistical learning and language processing

It is widely accepted that statistical learning is important for language acquisition and processing. For instance, statistical learning mechanisms are thought to be important for word segmentation (Saffran, Aslin, & Newport, 1996), word learning (Graf Estes, Evans, Alibali, & Saffran, 2007; Mirman, Magnuson, Graf Estes, & Dixon, 2008), and the acquisition of syntax (Gomez & Gerkin, 2000; Ullman, 2004). Previous work has shown that knowledge of the statistical probabilities in language can enable a listener to better identify – and perhaps even implicitly predict – the next word that will be spoken (Miller, Heise, & Lichten, 1951; Rubenstein, 1973; cf., Bar, 2007). This use of top-down knowledge becomes especially apparent when the speech signal is perceptually degraded, which is the case in many real-world situations. When ambient noise or multitalker babble degrades parts of a spoken utterance, the listener must rely on long-term knowledge of the statistical regularities in language to implicitly predict the next word that will be spoken based on the previous spoken words, thus improving speech perception and language comprehension (Elliott, 1995; Kalikow, Stevens, & Elliott, 1977; McClelland, Mirman, & Holt, 2006; Miller, et al., 1951; Pisoni, 1996).

Surprisingly, despite the voluminous work on statistical learning and the suggestions of its importance for language acquisition and processing, up until recently no studies had shown an empirical association between individual differences in statistical learning abilities and language. Recently, we investigated whether statistical learning abilities would be associated with one particular measure of everyday language performance: how well one uses preceding sentence context to implicitly predict upcoming linguistic units (Conway et al., 2010). The rationale is that statistical learning might provide a language user with knowledge that constrains the possible set of words that will be heard next in a sentence.

For example, consider the following two sentences:

(1) Her entry should win first prize.

(2) The arm is riding on the beach.

The final word in sentence (1) is highly predictable, while the final word in sentence (2) is not predictable. Therefore, when these two sentences are presented to participants under degraded listening conditions, long-term knowledge of language structure can improve perception of the final word in sentence (1) more so than in (2). We argue then, that performance on the first type of sentence ought to be more closely associated with fundamental statistical learning abilities because it relies on one's knowledge of word predictability that accrued implicitly over many years of exposure to language. On the other hand, performance on the second type of sentence simply relates to how well one perceives speech in noise, where knowledge of sequential word predictability is less useful.

We directly tested this hypothesis by assessing healthy adult participants on both statistical learning and speech perception tasks. In the statistical learning task, participants observed and then immediately reproduced visual color sequences on a touch-screen monitor (Figure 1). Unbeknownst to participants, the task consisted of two parts, a learning phase and a test phase, which differed only in terms of the sequences used. In the learning phase, the sequences were constrained such that only certain colors (e.g., blue) would ever occur following certain others (e.g., green). In the test phase, participants were now presented with novel sequences that either contained the same statistical regularities as before or completely random sequences in which any color could occur no matter what preceded it (except that immediate color repetitions were not allowed).

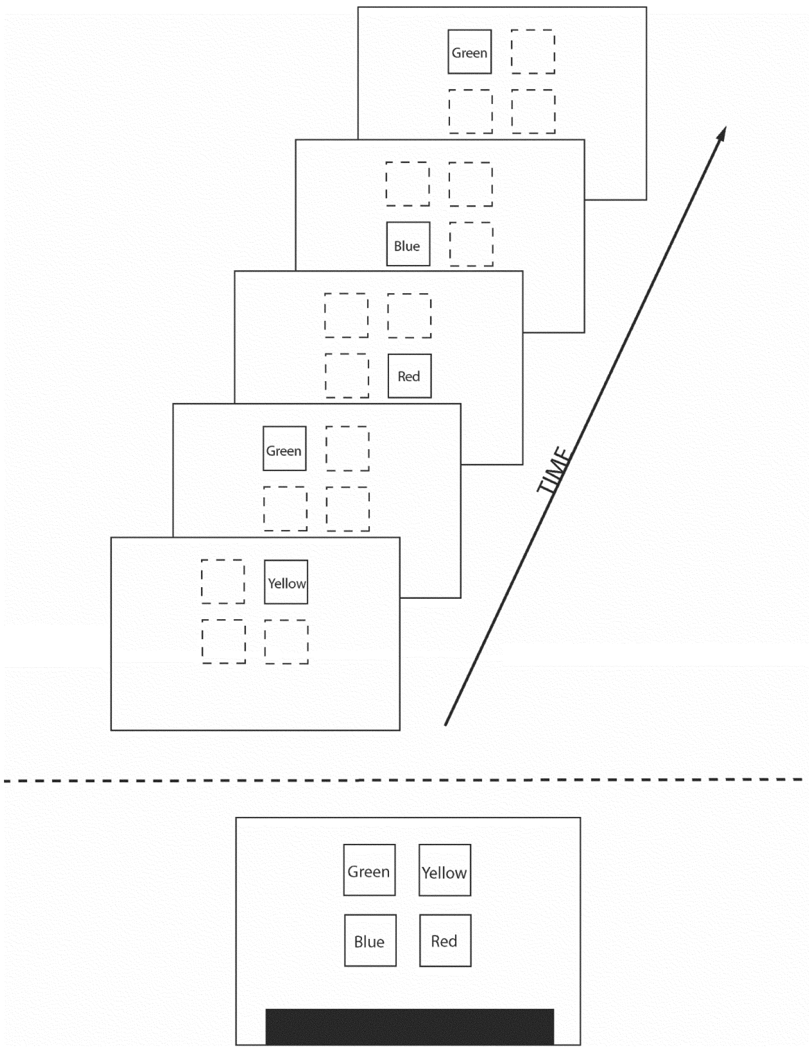

Figure 1. Depiction of the visual sequential statistical learning task used in Conway et al., 2010, similar to that used in previous work (Conway et al., 2007; Karpicke & Pisoni, 2004). Participants view a sequence of colored squares (700-msec duration, 500-msec ISI) appearing on the computer screen (top) and then, 2000-msec after sequence presentation, they must attempt to reproduce the sequence by pressing the touch-panels in correct order (bottom). The next sequence occurs 2000-msec following their response.

Learning was assessed by observing improvements to immediate memory span for statistically-consistent structured sequences (Botvinick, 2005; Conway et al., 2007; Hebb, 1961; Jamieson & Mewhort, 2005; Karpicke & Pisoni, 2004; Miller & Selfridge, 1950). That is, as participants were exposed to the statistical patterns, if any learning occurred, their immediate serial recall should improve for sequences that contained the same statistical regularities compared to ones that did not contain those regularities. This is an indirect measure of learning, which has a number of advantages over using a more traditional direct measure of learning such as explicitly asking for which sequence was more "familiar" or "obeyed the rules" (Redington & Chater, 2002). Importantly, this indirect measure appears to provide a wider range of individual differences in performance as compared to explicit measures of implicit and statistical learning (Karpicke & Pisoni, 2004).

Participants also completed a speech perception in noise task. In this task, participants had to identify sentences spoken under degraded listening conditions in which half of the sentences ended on a highly predictable word (sentences of type 1) and half ended on a low predictable word (sentences of type 2) (Elliott, 1995; Kalikow et al., 1977). To assess performance, we used the difference score suggested by Bilger and Rabinowitz (1979). This score was calculated by taking the difference between how well one perceives the final word in high-predictability sentences and how well they perceive the final word in low- or zero-predictability sentences. This difference score provides a means of assessing how well an individual is able to use sentence context to guide spoken language perception under degraded listening conditions.

Across three experiments, we found that individual differences in statistical learning abilities were significantly correlated with the sentence perception difference score (Figure 2). Importantly, the correlations remained even after controlling for sources of variance associated with non-verbal intelligence, verbal short-term memory and WM, attention and inhibition, and knowledge of vocabulary and syntax. We conclude that the common factor involved in both tasks – and which mediated the observed correlations – is sensitivity to the underlying statistical structure contained in sequential patterns, independent of general memory, intelligence, or linguistic abilities (Conway et al., 2010).

We propose that superior statistical learning abilities result in more detailed and robust representations of the structure of spoken language. Having a more detailed veridical representation of the likely probability that any given linguistic unit will follow based on what has already occurred

(V9 5/4/12 14:41) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 309)

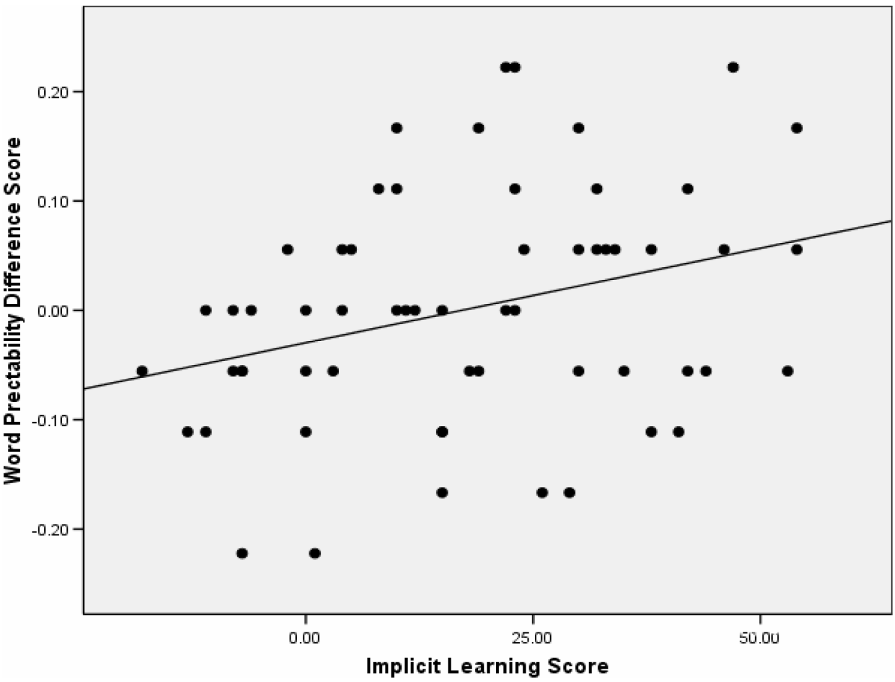

Figure 2. Scatterplot of data from Experiment 3 (n = 59) of Conway et al. (2010). The x-axis displays the implicit statistical learning scores; the y-axis displays the word predictability difference scores for the spoken sentence perception task.

can in turn improve how well one can rely on top-down knowledge to help implicitly predict, and therefore perceive, the next word spoken in a sentence. Thus, forming predictions for upcoming language units may be an important way in which statistical learning directly supports language acquisition and processing (see also Misyak, Christiansen, & Tomblin, 2010).

2. Statistical learning in language disorders

If domain-general statistical learning abilities are important for language acquisition and processing, then we might expect that what initially appear to be language-specific disorders may be due in part to disturbances with domain-general learning abilities. There is in fact, a growing body of evidence suggesting that this is indeed the case. Here, we review research examining statistical learning in specific language impairment and dyslexia. Then, we present the results of a new study with deaf children with cochlear implants that supports the theory that statistical learning is a crucial part of typical language acquisition, and if it is disturbed or developmentally delayed, can impair successful language development (Conway, Pisoni, Anaya, Karpicke, & Henning, 2011).

A growing body of research has established implicit learning impairments in individuals with various types of language disorders. For example, Plante, Gomez, and Gerken (2002) showed that a group of adults with language and reading impairments had more difficulty with an artificial grammar learning task than adults without a diagnosed language disorder. In terms of specific language impairment (SLI), recent research indicates that implicit learning abilities may be intact but significantly slower than in normal controls. For example, one study showed that in a serial reaction time task, adolescents with and without SLI showed evidence for implicitly learning embedded patterns (i.e., reaction times improved over trials), but learning rates for the SLI group were slower (Tomblin, Mainela-Arnold, & Zhang, 2007). Likewise, children diagnosed with SLI were able to learn an artificial language after 42 minutes, whereas controls learned it after only 21 minutes (Evans, Saffran, & Robe-Torres, 2009). It should be noted that, somewhat surprisingly, the SLI children were only able to learn the language when it was made up of speech (phonemic) stimuli; when it was made up of tones, performance for the SLI group did not reach above chance levels (Evans et al., 2009).

Regarding reading disorders such as dyslexia, the evidence on implicit learning is mixed. Studies using the visual serial reaction time task appear to show an absence of implicit learning. This is indicated by a failure to observe a decrease in reaction times to repeating patterns of stimuli by dyslexic participants (Menghini, Hagberg, Caltagirone, Petrosini, & Vicari, 2006; Vicari, Marotta, Menghini, Molinari, & Petrosini, 2003). However, other studies using techniques such as cued reaction time (Roodenrys & Dunn, 2008) and artificial grammar learning (Russeler, Gerth, & Munte, 2006) show unimpaired learning for individuals with dyslexia. Howard, Howard, Japikse, and Eden (1995) made a somewhat novel distinction: they showed that individuals diagnosed with dyslexia demonstrated normal learning on tasks involving spatial implicit learning, but showed impaired performance on tasks involving sequential implicit learning. With more research, this distinction may help resolve the previous divide in the literature on implicit learning in dyslexia.

A final population that offers an interesting test of the role of statistical learning in language is deaf children who have received a cochlear implant (CI). A CI is a medical prosthesis surgically implanted into the inner ear of a deaf child in order to provide sound by directly stimulating the auditory nerve. Although a CI provides the potential to develop age-appropriate speech and language abilities, it is well known that some children obtain little language benefit other than the awareness of sound from their implant (American Speech-Language-Hearing Association, 2004). Some of this variation in outcome has been shown to be due to certain demographic factors, such as age at implantation and length of deafness (Kirk et al., 2002; Tomblin, Barker, & Hubbs, 2007). However, these demographic variables leave a large amount of variance unexplained. It is likely that intrinsic cognitive factors, especially fundamental learning and memory abilities, contribute to language outcomes following implantation (Pisoni, 2000). Disturbances in statistical learning specifically may hold the key to understanding the enormous range of variation in language development in this population.

Deaf children with CIs also provide a unique opportunity to study neurocognitive plasticity and neural reorganization following the introduction of sound and spoken language after a period of auditory deprivation. Whereas most previous work with this clinical population has investigated the development of auditory perception, speech perception, and spoken language development following cochlear implantation, relatively few studies have examined more global learning and cognitive capabilities.

Recently we assessed visual sequential statistical learning abilities in a group of deaf children with CIs (Conway et al., 2011). Our aims were twofold: to assess the effects that a period of auditory deprivation and language delay may have on visual statistical learning skills; and to investigate the role that statistical learning plays in language outcomes following cochlear implantation. Our hypothesis was that deaf children with CIs would show disturbances in visual sequential statistical learning as a result of their relative lack of experience with sequential (auditory) patterns early on in development. Furthermore, we expected that statistical learning performance would be associated with measures of language development, with better statistical learners showing the best language outcomes post-implantation.

A group of deaf children with CIs engaged in a visual sequential learning task similar to the sequence reproduction task described earlier with adult participants. The results revealed that the CI children on average showed no learning (Figure 3, right), and were significantly worse than an age-matched group of hearing children (Figure 3, left).

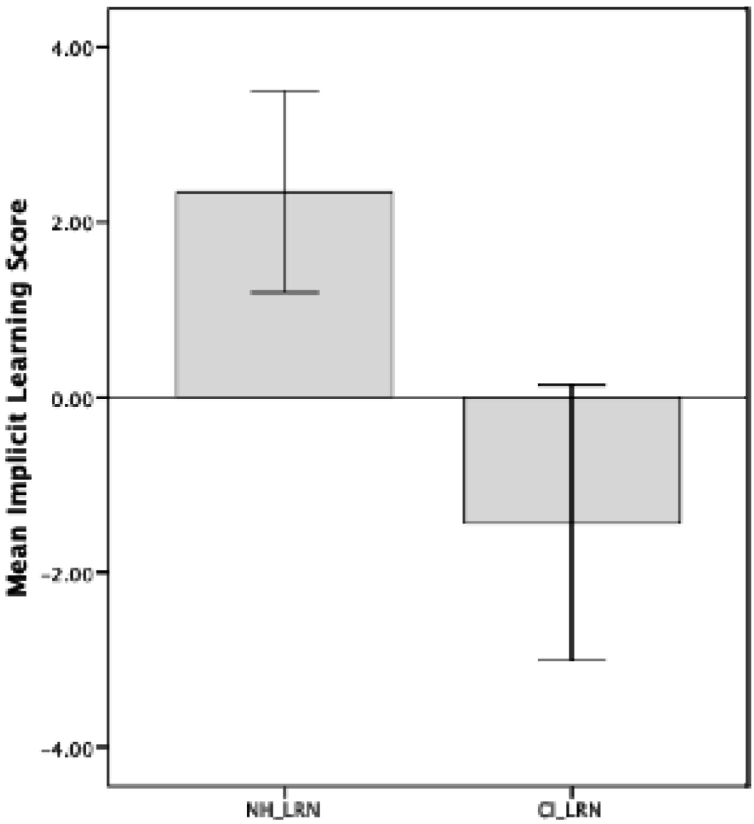

Figure 3. Average statistical learning scores for CI children (right) and NH children (left). Data adapted from Conway et al. (2011). Error bars represent ± 1 standard error.

Furthermore, performance on the statistical learning task was found to be significantly correlated with a standardized measure of language outcome, the Clinical Evaluation of Language Fundamentals, 4th Ed. (CELF-4; Semil, Wiig, & Secord, 2003), which has a particular emphasis on syntax-related language functions. That is, those children who were the best learners on the visual sequential statistical learning task showed the best language and syntax abilities as measured by the CELF-4. For the most part, these correlations remained significant even after controlling for the shared variance associated with duration of implant use (and age at which the device was implanted), forward and backward digit span, and vocabulary scores. In addition, performance on the statistical learning task was associated with how well the children could use sentence context to perceive spoken words (Pisoni, Conway, Kronenberger, Henning, & Anaya, 2010), a finding that is consistent with the adult data presented earlier.

Why did these children show a disturbance to non-linguistic visual statistical learning skills? There is some indication that a period of auditory deprivation occurring early in development may have secondary cognitive and neural ramifications in addition to the obvious hearing-related effects (Conway, Pisoni, & Kronenberger, 2009). Specifically, because sound is a temporally-arrayed signal, a lack of experience with sound may affect how well one is able to encode, process, and learn serial patterns (Marschark, 2006; Rileigh & Odom, 1972; Todman & Seedhouse, 1994). Exposure to sound may provide a kind of "auditory scaffolding" in which a child gains vital experience and practice with learning and encoding sequential patterns in the environment (Conway et al., 2009). We suggest that a lack of experience with sound may delay or alter the development of domain-general processing skills – including statistical learning – that rely on the encoding and learning of temporal or sequential patterns, even for non-auditory inputs. Poor sequential learning skills therefore might help explain why this particular population may have difficulty learning spoken language even after hearing is restored through a cochlear implant.

In sum, across a variety of populations having a language or communication disorder, we find that domain-general learning abilities are associated with the impairment, and therefore may provide a key for successful intervention and treatment through novel focused training techniques.

3. Study 1: Computerized training in healthy adults

The relationship between statistical learning and language in both healthy individuals and those with language disorders makes it important to ask whether it is possible to improve these domain-general learning abilities. A number of studies have demonstrated the efficacy of using different kinds of cognitive training paradigms to improve aspects of perception, attention, and cognition (Dye, Green, & Bavelier, 2009; Klingberg, in press; Rueda et al., 2005; Shalev, Tsal, & Mevorach, 2007; Tallal & Gaab, 2006).

To our knowledge, there have been no published attempts to improve statistical learning or any other non-declarative learning ability. However, one cognitive domain that has received much interest in the cognitive training literature is WM. While the training tasks and populations have varied, there is a growing body of evidence suggesting that computerized training tasks can improve WM capacity, and importantly, result in transfer to non-trained tasks of spatial and verbal WM, attention, and other cognitive functions (Curtis and D'Esposito, 2003, Olesen, Westerberg & Klingberg 2004, Holmes, Gathercole, & Dunning 2009, Thorell, Lindqvist, Nutley, Bohlin, & Klingberg 2009, Westerberg, Jacobaeus, Hirvikoski, Clevberger, Ostensson, Bartfai, & Klingberg 2007, Klingberg, Fernell, Olesen, Johnson, Gustafsson, Dahlstrom, Gillgberg, Forssberg, & Westerberg 2005; Verhaeghen et al., 2004).

The findings from these studies suggest that adaptive training on a visuospatial WM task appears to generalize in a domain-general manner to non-trained tasks of WM and other cognitive functions. For example, visuospatial WM training generalizes to inhibition (Klingberg et al. 2002; Klingberg et al. 2005, Oleson et al. 2004), attention (Westerberg et al. 2007), and verbal WM (Holmes et al. 2009; Thorell et al. 2009). Increased activity in the prefrontal cortex indicates that WM training has a direct impact on neural circuits of this brain region (Curtis & D'Esposito 2003; Olesen et al. 2004), implying that training tasks directly alter the functioning of domain-general executive control mechanisms (Smith & Jonides, 1999), rather than merely improving the efficiency of modality-specific "slave" systems. It has also been proposed that the striatum, a brain area recognized for its role in implicit learning (Seger, 2006), plays an important role in mediating transfer effects to non-trained tasks of WM (Dahlin, Bäckman, Neely, & Nyberg, 2009; Dahlin, Neely, Larsson, Bäckman, & Nyberg, 2008).

These studies demonstrate the utility of improving cognitive function through computerized training techniques, leaving open the possibility that like WM, statistical learning might also be amenable to training. As Klingberg (2010) rightfully pointed out, the synaptic mechanisms governing WM capacity are governed by the same principles of neural plasticity underlying the rest of the brain. Thus, we might expect that statistical learning can also be enhanced using a similar approach.

Whereas the standard short-term memory or WM task involves recalling a set of stimuli that have no intrinsic relation to each other, such as a series of random digits, most of our experiences in the world, such as events and scenes we encounter and interact with, have an underlying structure to them. How we acquire knowledge about these underlying regularities and statistical dependencies is arguably as important as how well we remember random, unstructured stimuli, if not more so. Enhancing statistical learning thus could have important and far-reaching ramifications, especially for populations having language delays or communication disorders.

(V9 5/4/12 14:41) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 315)

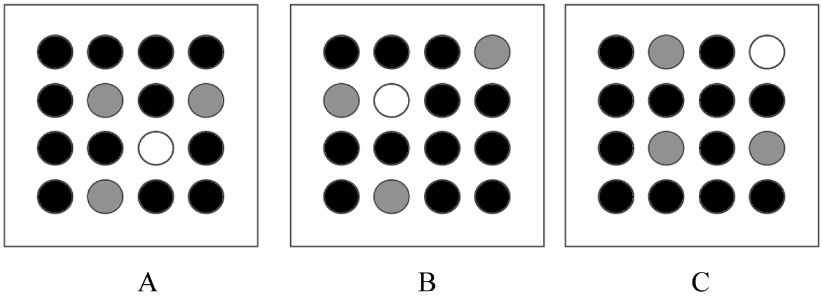

Figure 4. Computerized training task used in Study 1 and Study 2. Participants view a 4 × 4 matrix of circles. A sequence of circles light up, one at a time (the white circle depicted in each of the three scenes A, B, and C). Participants must reproduce each sequence in its entirety. Unknown to the participants, the presented sequences are not random; each circle can be followed by only 1 of 3 possible circles (shaded light grey). Note that in the actual task, all circles are colored the same, except for the one that is currently lit.

We recently created a novel computerized visual training task, and piloted it first with healthy adults (Bauernschmidt, Conway, & Pisoni, 2009; Conway, Bauernschmidt, & Pisoni, in preparation). The training task is a visual-spatial training procedure that is conceptually similar to other training tasks designed to improve WM abilities in adults and children (e.g., Holmes et al., 2009; Thorell et al., 2008). However, rather than using *random* sequences, we adaptively trained participants on *non-random* sequential patterns that share an underlying structure that can be implicitly learned. Thus, the novel facet of our task is that it adaptively trains participants to encode and reproduce sequences of visual stimuli conforming to underlying statistical regularities (see Figure 4).

In the training task, participants view a sequence of colored lights, occurring one at a time (white circles in panels A, B, and C) and then are required to reproduce what they saw by pressing the circles in correct order on the touch-sensitive monitor. Unbeknownst to the participants, each circle that lights up next in a sequence is not determined randomly but rather conforms to certain underlying statistical regularities. Specifically, any given circle has only three others that can legally follow it (shaded light grey). (Note that for the actual task, all circles are colored the same, except for the one that is currently lit.). As participants begin to implicitly uncover these regularities specifying which circles can occur

Table 1. Study 1 Training Schedule

Day 1 (Pre-Training)	Days 2–5 (Training)	Day 6 (Post-Training)
• Stroop Color and Word Test • Forward Digit Span • Backward Digit Span • Raven's Standard Progressive Matrices • Visual Sequential Learning	Group 1: Adaptive Training, with Statistically Constrained Sequences Group 2: Adaptive Training, with Pseudo-random Sequences Group 3: Non-Adaptive Control, using Pseudo-random Sequences	• Stroop Color and Word Test • Forward Digit Span • Backward Digit Span • Raven's Standard Progressive Matrices • Visual Sequential Learning

next, their performance on the recall task will improve, a sign of statistical learning occurring.

As with the WM training studies, an important aspect of this training task is that the lengths of the sequences presented to each participant are adaptively based on their performance level. Sequence lengths were based on a two-up, two-down metric. For example, if a subject starts at sequence length four and correctly reproduces all four items in that sequence, then their next trial will also be a sequence of length four. If the subject correctly reproduces all elements in the second sequence of length four, then they will move up to a sequence of length five in the next trial. If they *incorrectly* reproduce this sequence of length five then their next trial will still be sequence length five; if they respond incorrectly to this sequence as well, then their next sequence will be moved down to length four. And so on. Importantly, on each trial, a new sequence is presented (at the individual's current length). The new sequence is randomly determined, but conforms to the underlying regularities as specified earlier.

Participants engaged in this visual-spatial sequence training task for four days (days 2–5, see Table 1), with each training session lasting no longer than 45 minutes. Crucially, the "grammar" or statistical patterns that dictate what circles/locations can legally occur next were re-randomized for each participant on each subsequent training day. Because each of the four days of training incorporated a new set of statistical regularities, our intention was that participants would gradually improve their abilities to learn a variety of statistical patterns and not any one specific set of regularities. This was done to encourage generalization by improving participants' abilities to "learn to learn".

By adaptively training participants on this task, we expected to enhance their ability to learn statistical patterns of any type. The key test, of course, is whether any such improvements to learning result in transfer effects, that is, improvements to non-trained tasks. To test generalization and transfer, all participants were given a set of pre-training measures on Day 1 (see Table 1) that included the sequential statistical learning task used in Conway et al. (2010) in addition to other tests of verbal short-term memory and WM (measured with the Forward and Backward digit span tasks from the WISC-III, Wechsler, 1991), executive control and inhibition (measured by the Stroop Color and Word test, Golden & Freshwater, 2002), and nonverbal reasoning (measured by the Raven's Standard Progressive Matrices, Raven, Raven, & Court, 2000). These same measures were given once training was complete on Day 6 in order to ascertain improvements to these non-trained tasks.

Finally, in order to ensure that any observed gains on non-trained tasks were not merely a result of a test-retest effect, participants were randomly assigned to one of three different training conditions. Group 1 engaged in an adaptive, statistically-constrained version of the training task already described above. The Group 2 training task was identical to Group 1's except that the sequences were pseudo-random rather than conforming to statistical regularities. The pseudo-random sequences were generated so that each element (circle) in the sequence could be followed by any other in the set with equal likelihood. Like Group 1, the Group 2 task also was adaptive. Thus, Group 2 was very similar to previous WM training tasks. Finally, Group 3 served as a non-adaptive control using pseudo-random sequences. Participants in this group received visual sequences varying in length randomly determined at each trial, not based on their performance as was the case for Groups 1 and 2.

In sum, any training effects observed in Group 1 but not in Group 2 can be safely regarded as being due to the effect of including statistically-constrained sequences in the adaptive sequence task. Any training effects observed in Group 2 compared to Group 3 can be regarded as being due to the effect of using an adaptive (versus a non-adaptive) training paradigm.

Initial results are presented below for 56 adult participants (ages 18–30), with 20 participants in each of Groups 1 and 3, and 16 participants in Group 2 (Conway, Bauernschmidt, & Pisoni, in preparation). For each of the non-trained measures, a separate repeated measures ANOVA was run, with the within subject factor being the pre- vs. post-training dependent measure of interest, and the between subjects factor being training

(V9 5/4/12 14:41) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 318)

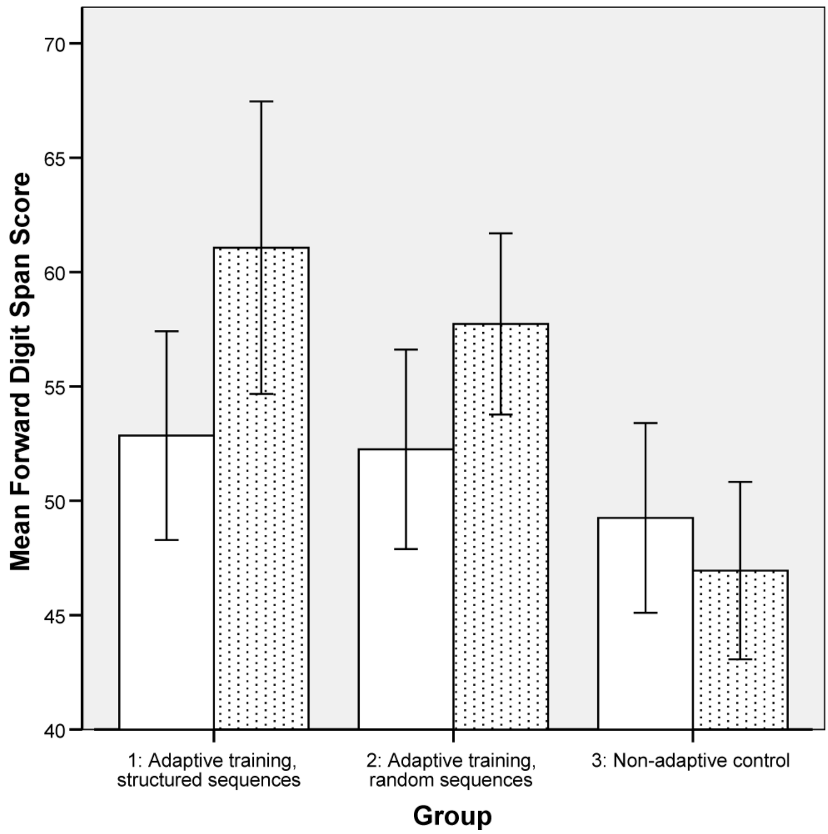

Figure 5. Effects of computerized training with healthy adults on Forward digit span scores (white bars: pre-training digit span scores; shaded bars: post-training digit span scores). Group 1 received adaptive training with statistically-constrained, structured sequences; Group 2 received adaptive training with pseudo-random sequences; and Group 3 served as a control group, receiving non-adaptive training with pseudo-random sequences. Error bars represent ± 1 standard error.

group (Groups 1–3). For all analyses, chronological age was used as a covariate. We report results from three of the non-trained measures below.

Figure 5 shows the pre- and post-training scores for each group on the Forward digit span task. In this task, which serves as a measure of verbal short-term memory capacity, subjects were presented with lists of spoken digits at progressive lengths and asked to repeat the sequence aloud. There was a marginally significant interaction of training group X pre- vs. post-scores: $F(2,45) = 2.62$, $p = .084$. Paired t-tests were used to compare pre-training to post-training performance for each of the three conditions, to determine for which training groups Forward digit spans improved (or worsened) following training. As shown in Figure 5, only Group 1 ($t(14) = 1.841$, $p = .087$) and Group 2 ($t(14) = 2.03$, $p = .062$) showed signs of improvement following training. Thus, adaptive training of visuospatial sequences showed signs of improving verbal auditory short-term memory capacity, regardless of whether the visuospatial sequences were statistically-constrained or pseudo-random. This transfer from a visuospatial to a verbal memory task is consistent with previous research showing training-related transfer across modalities (Thorell et al., 2008).

Next, Figure 6 shows pre- and post-training scores for each group on the Stroop Color and Word test. In this version of the classic task, participants were asked to read three pages of words, colors, and color-words aloud. The Word page consisted of the words "red", "green, and "blue" arranged randomly and printed in black ink. The Color page consisted of 100 items written as XXXX, printed in either red, green, or blue ink. The Color-Word page consisted of the words from the Word page printed in the colors from the Colors page. Participants are instructed to read the color of the print, not the word itself. Of course, for the Color-Word page, the words and colors do not always match, and as such, this requires inhibiting the natural and automatic response of reading the word.

There was a marginally significant interaction of training group X pre- vs. post-scores: $F(2,51) = 2.78$, $p = .07$. Similar to the Forward digit span results, performance on the Stroop task improved following training only for Group 1 ($t(18) = 3.04$, $p < .01$) and Group 2 ($t(15) = 1.86$, $p = .083$). The control Group 3 showed no signs of change. These results suggest that adaptive training of visuospatial sequences (statistically-constrained or pseudo-random) can improve executive control and inhibition abilities, also consistent with previous research (e.g., Klingberg et al., 2005).

Finally, and of most relevance at present, Figure 7 shows the pre- and post-training scores on a non-trained task of implicit statistical learning, the sequence learning task described earlier and used in several published studies (Conway et al., 2007; Conway et al., 2010). In this task, participants saw a sequence of four colored squares light up on the screen and were asked to reproduce the sequence that they had just seen by pressing the colors on a touch-screen monitor in correct order. Unbeknownst to the participants, the sequences were generated by an artificial grammar that provides statistical constraints on which color can occur next. Learning was assessed by computing a difference score for performance on statistically-

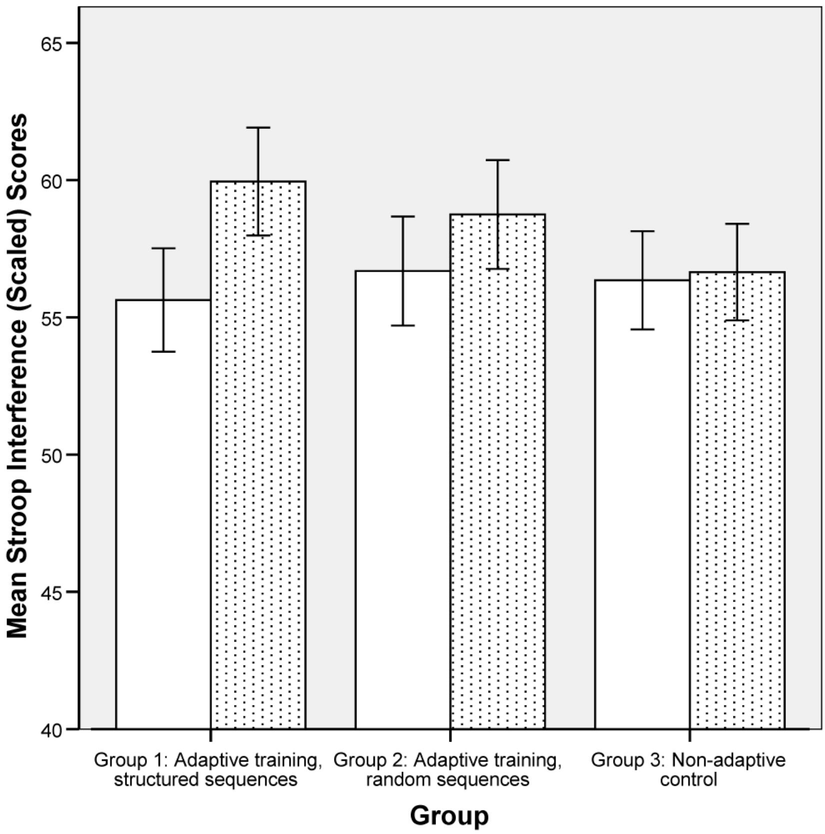

Figure 6. Effects of computerized training with healthy adults on Stroop interference scaled scores (white bars: pre-training Stroop scores; shaded bars: post-training Stroop scores). Group 1 received adaptive training with statistically-constrained, structured sequences; Group 2 received adaptive training with pseudo-random sequences; and Group 3 served as a control group, receiving non-adaptive training with pseudo-random sequences. Error bars represent ± 1 standard error.

constrained sequences minus performance on the sequences not conforming to the grammar.

There was a significant interaction of training group X pre- vs. post-scores: $F(2,50) = .71$, $p < .05$. Interestingly, for this non-trained sequential learning task, only Group 1 showed any hint of improvement ($t(19) = 1.37$, $p = .18$). Group 2 actually showed *worse* performance following training ($t(14) = 3.31$, $p < .01$); Group 3 showed no effect of training

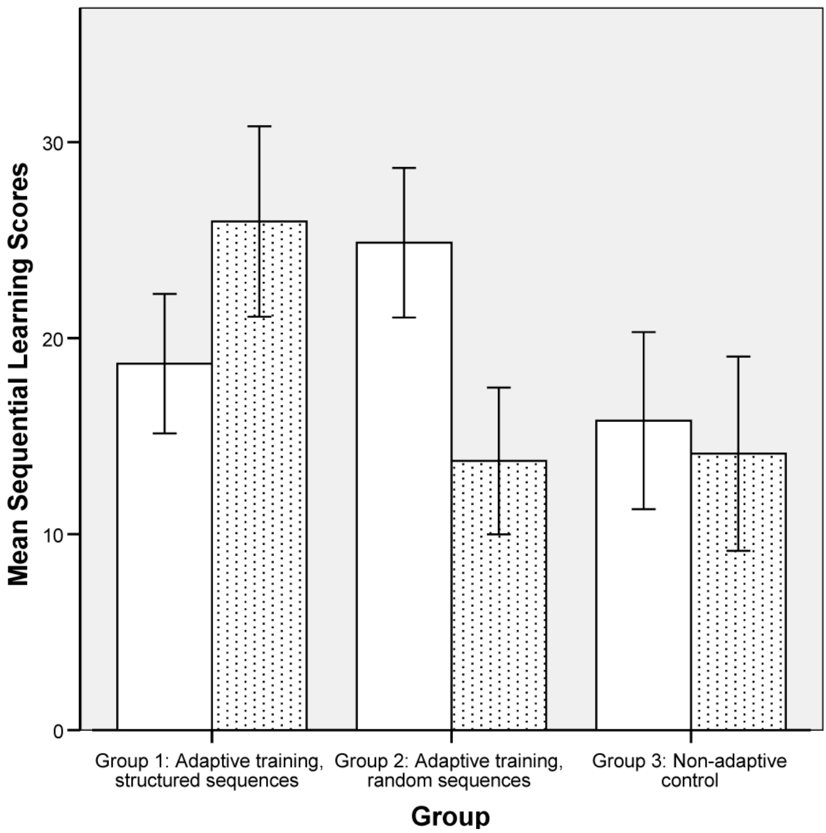

Figure 7. Effects of computerized training with healthy adults on a non-trained sequential learning task (white bars: pre-training sequential learning scores; shaded bars: post-training sequential learning scores). Group 1 received adaptive training with statistically-constrained, structured sequences; Group 2 received adaptive training with pseudo-random sequences; and Group 3 served as a control group, receiving non-adaptive training with pseudo-random sequences. Error bars represent ± 1 standard error.

either way. Thus, on this statistical learning measure we see a differential effect of using statistically-constrained versus pseudo-random sequences, with only the former showing any signs of improving statistical learning abilities on non-trained tasks.

In sum, these results suggest the following. First, only the adaptive training conditions (Groups 1 and 2) showed transfer effects to Forward digit spans and Stroop inhibition scores. This result is consistent with other recent findings demonstrating the utility of using computerized adaptive training to improve aspects of WM and executive function (Klingberg, 2010). Second, only the group that was specifically trained on sequential patterns with statistical structure (Group 1) showed any sign of improving on the non-trained sequential statistical learning task. In fact, the training condition that incorporated random sequential patterns actually led to significantly worse performance on the statistical learning task following training.

Although preliminary, these results suggest that training participants to interact with random patterns actually hampers their ability to learn structured patterns following training. On the other hand, training participants to interact with structured patterns not only leads to marginally better abilities to learn structured patterns following training, but also improves other WM and executive functions. Thus, incorporating statistically-structured patterns into a WM training task appears to provide just as much benefit as using unstructured random patterns and may actually show some carryover and transfer to other tasks requiring statistical learning. These findings provide initial evidence for the feasibility of improving domain-general learning abilities in populations that have a language delay, an endeavor we turn to next.

4. Study 2: Computerized training with deaf or hard of hearing children

As previously mentioned, there is now some evidence linking poor statistical learning abilities to impaired language function. Based on the findings in the previous section suggesting the possibility of improving statistical learning, we recently used our computerized training task with a group of children who are deaf or hard of hearing (d/hh) and who exhibit language delays to determine whether enhancing domain-general learning abilities can lead to improvements to overall language functioning. On a related note, Kronenberger et al. (in press) recently established the efficacy of using computerized WM training tasks to improve verbal and nonverbal WM in deaf children with cochlear implants, with some effects lasting up to 6 months.

In the present study, which is still ongoing, 23 children who are d/hh (ages 5:10 to 11:4; mean 8:2) took part in 10 days of training utilizing the computerized training task previously described. Among this group, 10 had bi-lateral cochlear implants, 8 were fitted with one implant and one

(V9 5/4/12 14:41) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 323)

hearing aid, and the remaining 5 children wore hearing aids in both ears. The children were assigned to one of two training conditions matched for chronological age. As with the adult study, sequences in the adaptive condition conformed to underlying statistical regularities, beginning at a length of three and increasing or decreasing in length based upon the two up, two down metric. The second condition was an active control group in which the sequence presentation was non-adaptive and pseudo-random in nature, with a constant sequence length of three. Pre- and post-training measures were selected to assess visual pattern memory, attention/inhibition, verbal WM, and visual sequential learning.

The children showed significant improvement on a number of non-trained tasks following training. Here, we focus on two of the measures, verbal WM and visual sequential learning. For the verbal WM task, a subset of 20 nonwords from the Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996) was presented to participants via a loud speaker at a level of 70–75 dBSPL. Participants were asked to repeat what they heard; responses were recorded then scored for overall word accuracy and for syllable accuracy, that is whether the response contained the same number of syllables as the stimuli presented. As shown in Figure 8, only children in the adaptive training condition showed a significant reduction in the number of syllable errors from the pre- to post-training session, $F(1,11) = 10.170$, $p = .009$, and this improvement was also sustained at a second post-training session measured 4–6 weeks later, $F(1,11) = 7.301$, $p = .021$.

This differential effect of sequence training, with only the adaptive group showing improvement, is also evident with a non-trained measure of visual sequential learning, as assessed by a version of the learning task described earlier. Figure 9 shows performance on this task with statistically-constrained sequences assessed before training and after the second post-test. Only participants in the adaptive condition showed significant improvement from pre-training to the second post-training session on the number of correctly reproduced statistically-constrained sequences, $F(1,11) = 9.308$, $p = .011$. Although the improvement in performance on the statistically-constrained sequences may suggest an improvement of sequential learning abilities, performance also improved on a set of pseudo-random sequences on this same task, raising the possibility that learning itself was not improved, but merely visual serial recall or sequential memory. We are currently exploring the viability of these alternative explanations.

(V9 5/4/12 14:41) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 324)

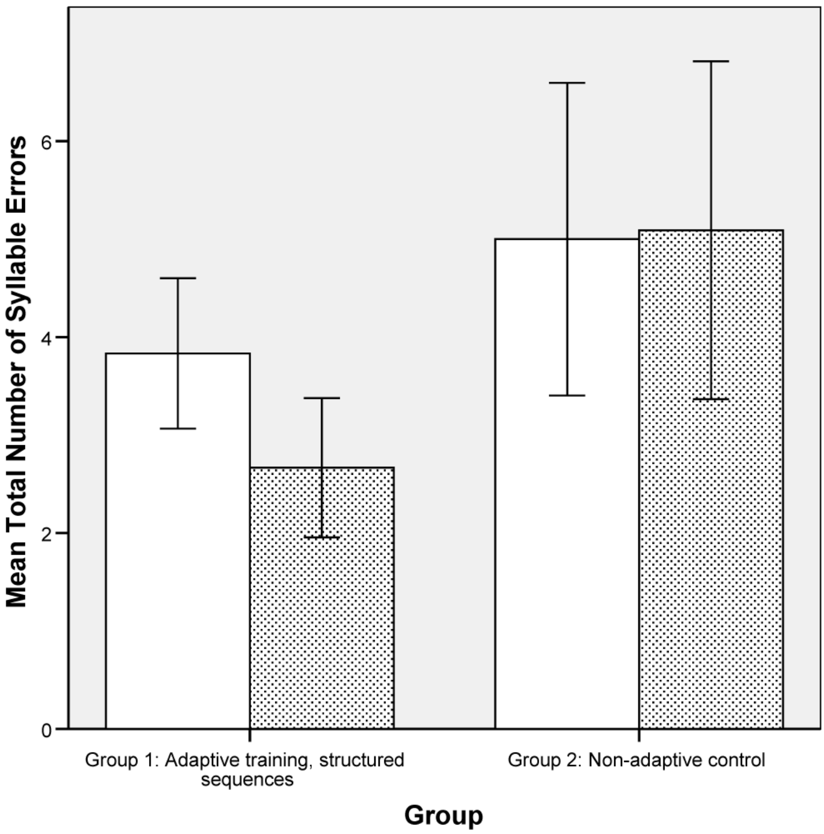

Figure 8. Effects of computerized training with deaf or hard of hearing children on nonword repetition errors (white bars: pre-training performance; shaded bars: post-training post-training performance). Group 1 received adaptive training with statistically-constrained, structured sequences; Group 2 served as a control group, receiving non-adaptive training with pseudo-random sequences. Error bars represent ±1 standard error.

In summary, these d/hh children showed significant improvement in the nonword repetition and non-trained sequential learning tasks. These findings hold promising implications for the improvement of language skills for children who are deaf or hard of hearing, and possibly other clinical populations as well. The nonword repetition task utilizes the same processes necessary for learning new words, namely, the completion of multiple auditory, cognitive, linguistic, and articulatory speech-motor processes, without the use of previous knowledge or visual cues. We would therefore expect

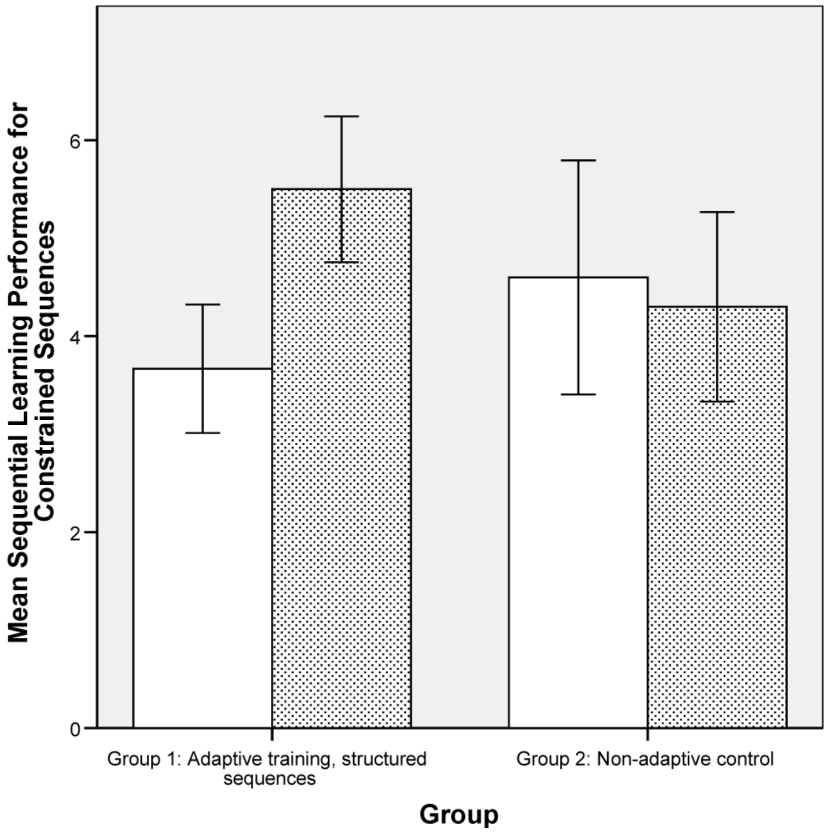

Figure 9. Effects of computerized training with deaf or hard of hearing children on a non-trained sequential learning task (white bars: pre-training performance; shaded bars: post-training performance). Group 1 received adaptive training with statistically-constrained, structured sequences; Group 2 served as a control group, receiving non-adaptive training with pseudo-random sequences. Error bars represent ± 1 standard error.

that improvement on this task would carry over to gains in receptive and expressive vocabulary and other language functions. So too, if statistical learning skills are improved as suggested by the current results, then it indicates that fundamental learning abilities important for language acquisition can be enhanced, and are likely to lead to more generalized global improvements to language, perhaps specifically for syntax-related aspects of language processing, a possibility that we will be exploring in the near future.

5. General discussion

Due to the increasing body of evidence suggesting that domain-general statistical learning abilities are used in the service of language acquisition, and given recent work showing the utility of using computerized adaptive training techniques, we believe it is important to attempt to improve statistical learning in order to treat language and communication disorders. The computerized training task that we have developed was based conceptually on recent WM training task designs. Our training task is relatively easy to implement and short in duration (45 minutes per day over 4–10 days) and crucially incorporates underlying statistical regularities into the patterns. The results with adults show that training resulted in gains to verbal short-term memory, executive control, and a non-trained task of sequential statistical learning. The findings from children who are deaf or hard of hearing also showed improvements to verbal short-term memory and sequential learning following training.

Although the findings are preliminary, they suggest that adaptive training of visuospatial statistically-constrained patterns can enhance broad domain-general skills of WM, inhibitory control, and statistical learning. This training task thus shows promise as a novel intervention for treating various disorders of language and learning.

If this training task does in fact improve performance on these three types of tasks (verbal short-term memory, inhibition, and statistical learning), it becomes important to ask what specific neurocognitive mechanism(s) were enhanced that led to these task improvements? Is there a common underlying function or set of functinos that are shared by all three tasks?

Fuster (2001) has argued that the prefrontal cortex (PFC) is critically involved in the temporal organization of behavior, including representing, formulating, and planning sequences of thought and action. For any complex sequential skill or behavior, the PFC is thought to be intimately involved because it allows for the integration of sensory cues with cognitive actions across time. Under this view, the PFC is important for any kind of sequencing or temporal functions (Conway & Pisoni, 2008), including higher level planning, executive memory, language processing, and sequential learning. The PFC has many interconnections with other sensory, motor, and subcortical regions, making it an ideal candidate for domain-general aspects of cognitive sequencing function (Miller & Cohen, 2001).

As other research has shown, adaptive WM training tasks appear to result in enhancements to the neural functioning of the prefrontal cortex, (Curtis & D'Esposito 2003; Olesen et al. 2004). For instance, Olesen et al.

(2004) had subjects practice three visuospatial memory tasks for a period of five weeks. Use of functional magnetic resonance imagining (fMRI) before, during, and after training showed increased activity in the prefrontal and parietal cortices. Similarly, Curtis and D'Esposito (2003) reported sustained prefrontal cortex activity during delay periods preceding the response portion of a visual WM task. The former study included a battery of neuropsychological tests as part of the pre- and post-training evaluation. Subjects showed significant improvement in performance on the Span board task and the Digit span task and in time on the Stroop test, illustrating, similarly to our results, transfer effects to non-trained tasks of WM and inhibition. These neuroimaging studies suggest that increases in cortical prefrontal activity during or following WM training is a sign of training-related plasticity in the neural systems supporting WM and other executive functions (Olesen et al. 2004).

Given the evidence of prefrontal activity and its relation to executive function, Funahashi (2001) proposed that the prefrontal cortex is responsible not only for storing and processing information, but also for assessing the input and providing information to neuronal systems to direct the processing of information in these systems (Funahashi, 2001). The processes of perception, motor control, and memory must be coordinated to accomplish the tasks of anticipating, planning, monitoring, and making a decision (Funahashi, 2001). The current evidence suggests that improvement on a visual-spatial sequence training task affects neural functioning of the prefrontal cortex and thus, perhaps by extension, executive and cognitive functions more generally, which may include statistical and sequential learning. The involvement of the prefrontal cortex in executive processes (Smith & Jonides 1999, Funahashi 2001) and evidence of increased prefrontal activity during spatial memory tasks (Curtis & D'Esposito 2003, Olesen et al. 2004, Smith & Jonides 1999, Funahashi 2001) thus lend support to the proposal that training on a visual-spatial task may carry over to other tasks involving different skills, including those requiring verbal memory or executive processing.

Although at present statistical learning is generally not considered to be an aspect of executive function, and if anything, might be rightfully thought of as a part of the nondeclarative/procedural learning system, there are reasons to believe that a connection may exist between some types of statistical learning and prefrontal cortical function. First, there is increasing neural evidence suggesting that the prefrontal cortex is involved during sequential learning and artificial grammar learning tasks (e.g., Fletcher et al., 1999; Forkstam et al., 2006; Petersson et al., 2004). Because our training task incorporates statistical patterns distributed across time – i.e., visual sequences – it is likely that the prefrontal cortex plays an important role in encoding these statistical regularities. Second, our training task likely promotes not merely statistical learning, but also cognitive control, attention, and inhibition. This is because at the beginning of every new training session, the "rules" or statistical regularities change, and so participants must over-ride the regularities that had been previously acquired. In this way, successful performance on this task requires participants to not only focus on the current input sequence, but to switch attention and inhibit prior learning in order to learn the new patterns. For these reasons, this training task may actually improve several overlapping elementary abilities (sequential learning, inhibition, attention, and serial recall) that are all mediated by the prefrontal cortex. Although we have no neural evidence yet, the behavioral evidence is consistent with this claim, with both verbal short-term memory and inhibitory control showing task gains following adaptive sequence training.

Of course, the ultimate objective remains to use these training tasks to improve learning and language abilities as a treatment for populations with language disorders. Notably, all three of the tasks that showed gains following training have been implicated as being important for language acquisition and processing: verbal short-term memory (Gathercole, Willis, & Baddeley, 1994), cognitive control (Deák, 2003), and of course, statistical and sequential learning (Conway et al., 2010). Clearly, the next step is to ascertain to what extent adaptive computerized training tasks such as this one that target statistical learning processes and other prefrontal cortex related abilities will show robust and lasting improvements to language function. In addition to treating language disorders, it may be possible to use this approach to help improve language acquisition for individuals learning a second language.

As this edited volume aptly indicates, we are beginning to realize the importance of domain-general statistical learning abilities for language. But we ought not to stop there. As recent research has amply demonstrated, our cognitive and neural systems are far more plastic and modifiable by experience than initially believed. By capitalizing on these theoretical and empirical developments, it may be possible to improve language functions by using novel computerized training techniques that specifically target domain-general learning abilities, offering great promise for alleviating disorders of language and communication.

(V9 5/4/12 14:42) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 329)

Acknowledgements

Portions of this work was supported by the following grants from the National Institutes of Health: R03DC009485, T32DC00012, R01DC00111, R01DC000064, and R01DC009581.

References

American Speech-Language-Hearing Association
2004 Cochlear implants [Technical report]. Available from www.asha.org/policy.

Bar, M.
2007 The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289.

Bauernschmidt, A., Conway, C.M., & Pisoni, D.B.
2009 Working memory training and implicit learning. In *Research on Spoken Language Processing Progress Report No. 29*. Bloomington, IN: Speech Research Laboratory, Indiana University.

Bilger, R.C., & Rabinowitz, W.M.
1979 Relationships between high- and low-probability SPIN scores. *The Journal of the Acoustical Society of America*, 65(S1), S99.

Botvinick, M.M.
2005 Effects of domain-specific knowledge on memory for serial order. *Cognition*, 97, 135–151.

Conway, C.M.
in press Sequential learning. Invited chapter to appear in R.M. Seel (Ed.), *Encyclopedia of the Sciences of Learning*. Springer Publications.

Conway, C.M., Bauernschmidt, A., Huang, S.S., & Pisoni, D.B.
2010 Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, 114, 356–371.

Conway, C.M., Karpicke, J., & Pisoni, D.B.
2007 Contribution of implicit sequence learning to spoken language processing: Some preliminary findings with hearing adults. *Journal of Deaf Studies and Deaf Education*, 12, 317–334.

Conway, C.M. & Pisoni, D.B.
2008 Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of New York Academy of Sciences*, 1145, 113–131.

Conway, C.M., Pisoni, D.B., Anaya, E.M., Karpicke, J. & Henning, S.C.
2011 Implicit sequence learning in deaf children with cochlear implants. *Developmental Science*, 14, 69–82.

Conway, C.M., Pisoni, D.B., & Kronenberger, W.G.
2009 The importance of sound for cognitive sequencing abilities. *Current Directions in Psychological Science*, 18, 275–279.

Curtis, C.E. & D'Esposito, M.
2003 Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7, 415–423.

Dahlin, E., Backman, L., Neely, A.S., & Nyberg, L.
2009 Training of the executive components of working memory: Subcortical areas mediate transfer effects. *Restorative Neurology and Neuroscience*, 27, 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, 1510–1512.

Deák, G.O.
2003 The development of cognitive flexibility and language abilities. *Advances in Child Development and Behavior*, 31, 271–327.

Dye, M.W.G., Green, C.S. & Bavelier, D.
2009 Increasing speed of processing with action video games. *Current Directions in Psychological Science*, 18, 321–326.

Elliott, L. L.
1995 Verbal auditory closure and the speech perception in noise (SPIN) test. *Journal of Speech, Language, and Hearing Research*, 38(6), 1363–1376.

Evans, J.L., Saffran, J.R., & Robe-Torres, K.
2009 Statistical learning in children with specific lanugage impairment. *Journal of Speech, Language, and Hearing Research*, 52, 321–335.

Fletcher, P., Buchel, C., Josephs, O., Friston, K. & Dolan, R.
1999 Learning related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cerebral Cortex*, 9, 168–178.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M. & Petersson, K.M.
2006 Neural correlates of artifial syntactic structure classification. *NeuroImage*, 32, 956–967.

Funahashi, S.
2001 Neuronal mechanisms of executive control by the prefrontal cortex. *Neuroscience Research*, 39, 147–165.

Fuster, J.
2001 The prefrontal cortex – An update: Time is of the essence. *Neuron*, 30, 319–333.

Gathercole, S.E. & Baddeley, A.D.
1994 *The Children's Test of Nonword Repetition*. London: Psychological Corporation Europe.

Gathercole, S.E., Willis, C.S., & Baddeley, A.D.
1994 The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory*, 2, 103–127.

(V9 5/4/12 14:42) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 331)

Gervain, J., & Mehler, J.
2010 Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, 61, 191–218.

Gogate, L.J., & Hollich, G.
2010 Invariance detection within an interactive system: A perceptual gateway to language development. *Psychological Review*, 117, 496–516.

Golden, C., & Freshwater, S.
2002 *The Stroop Color and Word Test*. Wood Dale, IL: Stoelting Co.

Gomez, R.L., & Gerken, L.
1999 Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109–135.

Graf Estes, K., Evans, J.L., Alibali, M.W. & Saffran, J.R.
2007 Can infants map meaning to newly segmented words? *Psychological Science*, 18, 254–260.

Gupta, P., & Dell, G.S.
1999 The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *The emergence of language* (pp. 447–481). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hebb, D.O.
1961 Distinctive features of learning in the higher animal. In J.F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–51). Oxford, Blackwell Scientific Publications.

Holmes, J., Gathercole, S. & Dunning, D.
2009 Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, F1–F7.

Howard Jr., J.H., Howard, D.V., Japikse, K.C., & Eden, G.F.
2006 Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia*, 44, 1131–1144.

Jamieson, R.K., & Mewhort, D.J.K.
2005 The influence of grammatical, local, and organizational redundancy on implicit learning: An analysis using information theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 31(1), 9–23.

Kalikow, D.N., Stevens, K.N., & Elliott, L.L.
1977 Development of a test of speech intelligibility in noise using materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61(5), 1377–1351.

Karpicke, J.D., & Pisoni, D.B.
2004 Using immediate memory span to measure implicit learning. *Memory & Cognition*, 32(6), 956–964.

Kirk, K.I., Miyamoto, R.T., Lento, C.L., Ying, E., O'Neil, T., & Fears, B.
2002 Effects of age at implantation in young children. *Annals of Otology, Rhinology, & Laryngology*, 189, 69–73.

Kleim, J.A., & Jones, T.A.
2008 Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*, 51, 225–239.

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

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

Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlstrom, 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 & Adolescent Psychiatry*, 44, 177–186.

Korkman, M. Kirk, U., & Kemp, S.
2007 *NEPSY-II*. San Antonio, TX: Psychological Corporation.

Kronenberger, W.G., Pisoni, D.B., Henning, S.C., Colson, B.G., & Hazzard, L.M.
in press Working memory training for children with cochlear implants. *Journal of Speech, Language, & Hearing Research*.

Kuhl, P.K.
2004 Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831–843.

Marschark, M.
2006 Intellectual functioning of deaf adults and children: Answers and question. *European Journal of Cognitive Psychology*, 18, 70–89.

McClelland, J.L., Mirman, D., & Holt, L.D.
2006 Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10, 363–369.

Menghini, D., Hagberg, G.E., Caltagirone, C., Petrosini, L., & Vicari, S.
2006 Implicit learning deficits in dyslexic adults: and fMRI study. *NeuroImage*, 33, 1218–1226.

Miller, G.A. & Cohen, J.
2001 An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.

Miller, G.A., Heise, G.A., & Lichten, W.
1951 The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41, 329–335.

Miller, G.A., & Selfridge, J.A.
1950 Verbal context and the recall of meaningful material. *American Journal of Psychology*, 63, 176–185.

Mirman, D., Magnuson, J.S., Estes, K.G., & Dixon, J.A.
2008 The link between statistical segmentation and word learning in adults. *Cognition*, 108, 271–280.

(V9 5/4/12 14:42) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 333)

Misyak, J.B., Christiansen, M.H., & Tomblin, J.B.
2010 Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science*, 2, 138–153.

Nicolson, R.I., & Fawcett, A.J.
2007 Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neuroscience*, 30(4), 135–141.

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

Perruchet, P. & Pacton, S.
2006 Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5), 233–238.

Petersson, K.M., Forkstam, C. & Ingvar, M.
2004 Artificial syntactic violations activate Broca's region. *Cognitive Science*, 28, 383–407.

Pisoni, D.B.
1996 Word identification in noise. *Language and Cognitive Processes*, 11(6), 681–687.

Pisoni, D.B.
2000 Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear & Hearing*, 21, 70–78.

Pisoni, D.B., Conway, C.M., Kronenberger, W., Henning, S., & Anaya, E.
2010 Executive function, cognitive control, and sequence learning in deaf children with cochlear implants. In M. Marschark & P. Spencer (Eds), *Oxford Handbook of Deaf Studies, Language, and Education* (pp. 439–457). New York, NY: Oxford University Press.

Plante, E., Gomez, R., & Gerken, L.
2002 Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*, 35, 453–462.

Raven, J., Rave, J.C., & Court, J.
2000 Standard Progressive Matrices. San Antonio, TX: Harcourt Assessment.

Reber, A.S.
1967 Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.

Redington, M., & Chater, N.
1997 Probabilistic and distributional approaches to language acquisition. *Trends in Cognitive Sciences*, 1(7), 273–281.

Redington, M., & Chater, N.
2002 Knowledge representation and transfer in artificial grammar learning (AGL). In R.M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, philosophical, and com-*

putational consensus in the making (pp. 121–143). Hove, East Sussex: Psychology Press.

Rileigh, K.K. & Odom P.B.
1972 Perception of rhythm by subjects with normal and deficient hearing. *Developmental Psychology*, 7, 54–61.

Roodenrys, S., & Dunn, N.
2008 Unimpaired implicit learning in children with development dyslexia. *Dyslexia*, 14, 1–15.

Rubenstein, H.
1973 Language and probability. In G. A. Miller (Ed.), *Communication, language, and meaning: Psychological perspectives* (pp. 185–195). New York: Basic Books.

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*, 102, 14931–14936.

Russeler, J., Gerth, I., & Munte, T.F.
2006 Implicit learning is intact in adult developmental dyslexic readers: Evidence from serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology*, 28, 808–827.

Saffran, J.R.
2003 Statistical learning learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12(4), 110–114.

Saffran, J.R., Aslin, R.N., & Newport, E.L.
1996 Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.

Seger, C.A.
2006 The basal ganglia in human learning. *The Neuroscientist*, 12, 285–290.

Semel, E., Wiig, E.H., & Secord, W.A.
2003 *Clinical evaluation of language fundamentals, fourth edition (CELF-4)*. Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Shalev, L., Tsal., Y. & Mevorach, C.
2007 Computerized progressive attentional training (CPAT) program: Effective direct intervention for children with ADHD. *Child Neuropsychology*, 13, 382–388.

Smith, E.E. & Jonides, J.
1999 Storage and executive processes in the frontal lobes. *Science*, 283, 1657–1661.

Tallal, P. & Gaab, N.
2006 Dynamic auditory processing, musical experience, and language development. *Trends in Neuroscience*, 29, 382–390.

(V9 5/4/12 14:42) WDG-LCB (155mm×230mm) TimesNRMT 1382 Rebuschat pp. 305–336 1382 Rebuschat_10_conway (p. 335)

Thorell, L., Lindqvist, S., Nutley, S., Bohlin, G. & Klingberg, T.
2009 Training and transfer effects of executive functions in preschool children. *Developmental Science*, 1–8.

Todman, J. & Seedhouse, E.
1994 Visual-action code processing by deaf and hearing children. *Language and Cognitive Processes*, 9, 129–141.

Tomblin, J.B., Barker, B.A., & Hubbs, S.
2007 Developmental constraints on language development in children with cochlear implants. *International Journal of Audiology*, 46, 512–523.

Tomblin, J.B., Mainela-Arnold, E., & Zhang, X.
2007 Procedural learning in adolescents with and without specific language impairments. *Language Learning and Development*, 3, 269–293.

Ullman, M.T.
2004 Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92, 231–270.

Ullman, M.T., & Pierpont, E.I.
2005 Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, 41, 399–433.

van Praag, H., Kempermann, G., & Gage, F.H.
2000 Neural consequences of environmental enrichment. *Nature Reviews: Neuroscience*, 1, 191–198.

Verhaeghen, P., Carella, J. & Basak, C.
2004 A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30, 1322–1337.

Vicari, S., Marotta, L., Menghini, D., Molinari, M., & Petrosini, L.
2003 Implicit learning deficit in children with developmental dyslexia. *Neuropsychologia*, 41, 108–114.

Wechsler, D.
1991 *Wechsler Intelligence Scale for Children – Third Edition*. San Antonio, TX: The Psychological Corporation.

Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Ostensson, M.L., Bartfai, A. & Klingberg, T.
2007 Computerized working memory training after stroke – A pilot study.