Training and plasticity of working memory

Torkel Klingberg

Department of Neuroscience, Karolinska Institute, Retzius väg 8, 171 77 Stockholm, Sweden

Working memory (WM) capacity predicts performance in a wide range of cognitive tasks. Although WM capacity has been viewed as a constant trait, recent studies suggest that it can be improved by adaptive and extended training. This training is associated with changes in brain activity in frontal and parietal cortex and basal ganglia, as well as changes in dopamine receptor density. Transfer of the training effects to non-trained WM tasks is consistent with the notion of training-induced plasticity in a common neural network for WM. The observed training effects suggest that WM training could be used as a remediating intervention for individuals for whom low WM capacity is a limiting factor for academic performance or in everyday life.

Explicit versus implicit training of working memory

Working memory (WM) refers to the retention of information over a brief period of time, a function that is of central importance for a wide range of cognitive tasks and for academic achievement [1]. Impaired WM is observed in many neuropsychiatric conditions, such as traumatic brain injury, stroke, mental retardation, schizophrenia and attention-deficit hyperactivity disorder (ADHD) [2]. It is thus not surprising that attempts to improve WM have a long history. In their 1972 article 'On the theory and practice of improving short-term memory', Earl Butterfield and colleagues reported a series of studies attempting to improve short-term memory in learning-disabled individuals by teaching participants to use subvocal rehearsal strategies [3]. Although this approach led to some improvement in participants' performance, there was no evidence of transfer either to non-trained cognitive tasks or to everyday performance.

A later study by Ericsson and colleagues demonstrated that the number of digits an individual can remember can be greatly improved by practice [4]. In particular, participant S.F. could after practice recall a series of 79 digits after hearing it only once by grouping numbers and associating them with running times stored in long-term memory. This strategy is an example of 'chunking', whereby isolated pieces of information are put together to form a meaningful combination that can be associated with previously stored long-term memories. However, this type of enhancement is highly material-specific, as demonstrated by the fact that when tested on his ability to remember letters, S.F. could only recall six. His WM capacity, defined as a trait that influences performance in WM tasks irrespective of the material to be memorized, was thus not improved.

These early training studies seemed to provide support for the static view of WM formulated by Miller in his article 'The magic number seven' [5]. However, subsequent research has shown that training can improve performance in a wide range of functions and that this improved performance is associated with neuronal changes from the intracellular level to functional organization of the cortex [6]. Training on motor [7] and perceptual tasks [8] in animals leads after hundreds of trials to enhanced performance, with concomitant changes in synaptic connectivity in both sensory and motor areas. Plasticity has also been demonstrated in the prefrontal cortex in animals [9].

This type of perceptual and motor training might be called implicit because improvement is based only on repetition, feedback and often gradual adjustment of the

Glossary of neuropsychological tests

Block design task: the subject is asked to arrange blocks that are either red or white or both red and white so that the blocks form a specific pattern [67]. The test measures spatial perception and problem solving.

Bochumer Matrizen test (BOMAT): this test is similar to Raven's matrices but was developed more recently [68].

Continuous performance task (CPT): a continuous series of digit or letters is presented either visually or auditorily [43]. The participant is asked to respond (e.g. by pressing a button) when a specific combination of stimuli, such as an A followed by an X, is presented. The task was developed to measure sustained attention and vigilance.

Paced auditory serial addition task (PASAT): in this task, participants are presented with a continuous stream of digits and asked to add the two final digits and report the sum aloud [69]. The task is considered a measure of both WM and sustained attention.

Span board task: this task measures visuospatial WM. Ten blocks are positioned on a brick. The tester points to a sequence of blocks and the participant is asked to reproduce the sequence [70].

Stroop task: the color words 'blue',' green',' red' and 'yellow' are randomly presented. In one version of the task the color in which words are printed is congruent with the meaning of the words (e.g. GREEN, RED, etc.) [42]. In another version the colors are incongruent (e.g. GREEN, RED) and subjects are asked to name the color in which the word is printed. The test measures a subject's ability to inhibit the prepotent response of reading the words.

Raven's matrices: this test measures inductive reasoning ability. A matrix of figures is presented in which one position is empty. By deducing the relationship between rows and columns, the participant is required to infer what figure should be in the empty position of the matrix [71]. Several versions of Raven's matrices exist, including the colored progressive matrices, standard progressive matrices and advanced progressive matrices.

Wechsler abbreviated scale of intelligence (WASI): this battery comprises tests of vocabulary, judgment of similarities, block design and matrix reasoning tasks [72].

Corresponding author: Klingberg, T. (torkel.klingberg@ki.se).

1364-6613/$ – see front matter doi:10.1016/j.tics.2010.05.002 Trends in Cognitive Sciences 14 (2010) 317–324

difficulty. By contrast, teaching of strategies to improve performance in WM tasks, such as rehearsal [3], chunking and meta-cognitive strategies [10], are explicit in that they are conscious strategies for handling the material. The question is whether implicit WM training might lead to durable neuronal changes in WM-related areas in the same way as perceptual training does for neurons of the visual cortex. The hypothesis put forward in the present article is that there is nothing magical about WM: the synaptic connections determining WM capacity are governed by the same laws of plasticity that characterize other parts of the brain.

Psychological and neural correlates of WM

Neurophysiological studies show that maintenance of information in WM is associated with elevated and sustained neural firing over a delay when information is kept in mind [11]. Neuroimaging studies in humans have mapped WM-related activity to both sensory association cortices and prefrontal cortex [12,13]. Some of these regions show specificity to the sensory modality of the stimuli [12,13]. Other regions, including parts of the intraparietal cortex and dorsolateral prefrontal cortex, are activated across several modalities and thus reflect a multimodal type of activity [12–14]. Individual differences in activity in both the intraparietal and prefrontal cortex are correlated to WM capacity differences among adults [15–19], as well as when comparing children and adults [20–25]. Neural network models have suggested that stronger frontoparietal connectivity is one potential mechanism behind higher WM capacity [26,27].

Mapping of the neural activity during WM tasks to specific psychological terms is still work in progress. Psychological models of WM distinguish sensory-specific storage from a coordinating or controlling function, referred to as the central executive [28] or controlled attention [29]. Attention is thus closely linked to WM. Controlled, or top-down, attention refers to the voluntary allocation of selective attention and relies on parietal and prefrontal regions that largely overlap with activation during WM tasks in both the parietal and prefrontal cortex [30]. Control of attention is necessary in WM tasks, for example when selecting only relevant information [17,31]. Conversely, retention of an internal representation of a salient location in WM is crucial for directing and maintaining attention towards that location [32]. The neural basis for WM and controlled attention might thus rely on the same mechanisms of sustained neural activity and top-down excitation, and the same multi-modal frontoparietal network, and might be difficult or impossible to separate even at the neuronal level.

The effect of training on a particular cortical region using a specific task would only be expected to transfer to other tasks and functions to the extent that the tasks rely on the same neural networks [33]. Training affecting sensory association areas would not be expected to have transfer effects to other modalities. Training affecting higher association cortices, however, might have more general effects. In particular, neural changes in the common intraparietal–prefrontal network would be expected to improve performance in WM tasks irrespective of the sensory modality of the memoranda, as well as in tasks requiring control of attention.

Computerized training of WM

An example of what might be termed implicit WM training is the training program originally developed by Klingberg and colleagues for children with ADHD [34,35]. This training involves repeated performance of WM tasks, with feedback and rewards based on the accuracy for every trial. The effective training time is 30–40 min per day, 5 days a week for 5 weeks (totaling approx. 15 h). The difficulty of the tasks is adjusted during the WM training on a trial-by-trial basis by changing the amount of information to be remembered so that it is close to the capacity of the subject.

This approach differs from previous WM training attempts in several ways. First, the training was not designed to teach explicit strategies, such as rehearsal techniques or meta-cognitive strategies [3,10]. Second, the method differs in the amount of time spent specifically on WM tasks. Previous interventions typically used WM tasks as part of training batteries that included other types of executive functions tasks, which decreased the overall time spent on WM tasks [36]. Third, the use of computerized tasks rather than typical one-on-one testing made it possible to have longer training times and to change the WM load on a trial-by-trial basis.

Results obtained using this particular method are briefly summarized before alternative, but also implicit, methods of WM training are reviewed (for methodological issues in training studies see Box 1).

Increase in WM capacity

Using the method developed by Klingberg et al., several studies have shown transfer improvement in WM tasks that were not part of the training program (Table 1) [34,35,37–40]. In several of these studies the active control

Box 1. Methodological aspects of the evaluation of cognitive interventions

The principles for studying cognitive training are the same as those for evaluating other medical interventions and pharmacological treatments. Some of the issues to consider include:

- **Including an active control group**. In pharmacological terminology, this is the placebo group. The active control group receives a believable alternative treatment to control for effects of expectancy, which is known to effect cognitive performance [78]. Computerized cognitive training also involves several non-specific aspects, such as adherence to a training schedule, visits to a clinic, use of a computer, etc. A passive control group or wait-list control group does not control for any of these effects.
- **Evaluating transfer**. Repeated performance of a task always leads to improved performance on that particular task. The important question, both theoretically and for the possible usefulness of training, is the extent to which the training can be generalized to non-trained tasks. The level of transfer can be difficult to quantify. It could be graded from: (i) transfer within the same domain (e.g. WM) but to other stimuli and a different response mode; (ii) transfer to other cognitive constructs (e.g. from WM to non-verbal reasoning); ideally, such effects should be evaluated by latent variables created from several tasks [79]; to (iii) transfer to everyday behavior. For clinical purposes, the ultimate goal is often to affect quality of life. Evaluation of the cognitive aspects of behavior is methodologically difficult. Rating scales are commonly used, but these are subjective and relatively coarse.

Table 1. Behavioral studies of WM training[a]

Ref.	Training	Population (age in years)	Control group	Transfer to WM	Transfer to other cognitive tasks	Transfer to reasoning
[34]	CWMT	ADHD (7–15)	Active	Span board*	Stroop*	Raven CPM*
[35]	CWMT	ADHD (7–12)	Active	Span board*	Stroop*	Raven CPM*
[39]	CWMT	Stroke patients (34–65)	Passive	Span board*	PASAT* Stroop, n.s.	Raven SPM, n.s.
[40]	CWMT, spatial	Healthy volunteers (4–5)	Active	Word span*	CPT* Go/no go, n.s.	Block design, n.s.
[37]	CWMT	Low WM (8–11)	Active	Complex span* Instruction WM*	Mathematical reasoning*	WASI, n.s.
[38]	CWMT	ADHD (8–11)	Repeated baseline	Complex span*		WASI, n.s.
[45]	*n*-back, lists	Healthy volunteers (20–70)	Passive	*n*-back* Digit span, n.s.	Recall of nouns*	Raven APM, n.s.
[46]	*n*-back	Healthy volunteers (20–80)	Passive	*n*-back* Complex span, n.s.		
[47]	*n*-back, dual	Healthy volunteers (mean age 26)	Passive	Digit span* Complex span, n.s.		BOMAT*

[a]CWMT, computerized WM training, as described by Klingberg *et al.* [35]; CPM, colored progressive matrices; SPM, standard progressive matrices; PASAT, paced auditory serial addition task; CPT, continuous performance task; WASI, Wechsler abbreviated scale of intelligence; APM, advanced progressive matrices; BOMAT, Bochumer Matrizen test.

*Significant change compared to control group; n.s., tested but not significant.

group performed non-adaptive WM training (repeated performance of WM tasks at a level far below the capacity limit), which provides a more stringent control than passive control groups [34,35,37].

Two studies reported transfer effects in many different types of WM tasks, including complex span tasks, which have the dual requirement of performing a task while keeping information in mind, even though no such tasks were included in the training program [37,38]. In another study, pre-school children performed WM training exclusively with visuospatial WM tasks, but showed significant improvement on a verbal WM task, thus clearly demonstrating generalization between modalities [40]. Transfer was also demonstrated by improvement in an ecologically relevant WM task called a ‘following instructions task’, which consists of remembering and performing instructions of increasing length (‘take the blue pen and put it in the red box’). In these studies (Table 1), improvements in tasks that were part of the training program were approximately 30–40%, whereas improvements in non-trained WM tasks were ~15%, with effect sizes (Cohen’s δ) relative to the control group of approximately 1.0 [35,37].

Taken together, these studies suggest that this WM training program [35] inspired by perceptual training methods leads to improvements in general WM capacity, as evidenced by improved performance in non-trained tasks varying in the type of material and mode of testing. The effect remained significant at 3-month [35] and 6-month [37] follow-up testing. These findings are thus consistent with the notion of training-induced plasticity in a common neural network for WM.

Improvement in inhibition, reasoning and inattentive behavior

Inhibitory functions and reasoning are closely related to WM [1,41]. Potential transfer effects of WM training to performance in non-WM tasks were investigated in several studies. Improvement in the Stroop task [42] was observed in two training studies of children with ADHD [34,35], as well as in young adults [33], but not in stroke patients [39]. However, it should be noted that control congruent trials were not included in these studies. It is thus not clear if the improvement observed is specific to inhibition or more generally to top-down attention. In a study of normal preschool children [40], trained children made fewer commission errors in a continuous performance task [43], but not in a go/no-go task. Improvements in reasoning tasks are weaker than for untrained WM tasks (effect size of 0.4) [35] and sometimes non-significant (Table 1) [37]. These inconsistencies could possibly be due to differences between populations and the specific tasks used to assess transfer effects.

WM training in children with ADHD led to a significant decrease in the number of inattentive ADHD symptoms in a controlled study that was evaluated by blinded raters [35] and in a pilot study of children with attentional problems from low socioeconomic status background [44]. A decrease in cognitive symptoms was also noted in a study of stroke patients [39]. This is consistent with the hypothesized overlap between neural mechanisms underlying the control of attention and those responsible for WM (see above).

WM training focusing on updating

The training program described above focused on training and increasing WM capacity, primarily by increasing the amount of visuospatial information that should be retained. Another approach to WM training also uses the principles of implicit training, but focuses specifically on updating, namely the replacement of old information in a hypothetical WM store with new information [45–47].

In a study by Dahlin *et al.*, young and old healthy adults practiced three computerized updating tasks for 45 min per session, three times a week for 5 weeks (11 h of training in total) [45]. During training, participants were presented with lists of letters, digits, colors or spatial locations, and at the end of the list recalled the last five items. Difficulty was adjusted by presenting longer lists but not changing the number of items to be recalled, that is, the load. Compared to a passive control group, the younger training group (but not the older group) improved significantly in a non-trained 3-back WM task, an effect that remained significant at 18-month follow-up testing. However, there was no improvement in three other WM tasks (forward or backward digit span and a computation span task) nor on Raven's advanced progressive matrices. Li and collaborators evaluated the effect of updating training in young and old adults, with 45 days of 15-min training (total 11 h) [46]. The training tasks were two versions of a spatial 2-back task. Both the young and old training groups improved in a spatial and a numerical 3-back task, but there were no improvements in two complex span tasks (a rotation span and a computation span task). In both studies the only significant transfer within the WM domain was to other WM tasks sharing the updating, or *n*-back mode of stimulus presentation, suggesting that the improvement was restricted to task-specific aspects of the updating tasks [45,46].

A slightly different approach was used in a study by Jaeggi *et al.* in which participants were trained on a dual *n*-back task with simultaneous matching of both positions and numbers to previously presented items [47]. Difficulty was adjusted by increasing the length of the lag, starting with 1-back and proceeding as high as 5-back at the end of training. Participants trained for 25 min per day for 8–19 sessions (maximum training time ~8 h). Compared to a passive control group, training improved performance in a digit span task and a test of non-verbal reasoning, the Bochumer Matrizen test (BOMAT). Significant improvement emerged between 12 and 17 days of training. The transfer effect on non-verbal reasoning is interesting and consistent with previous reports of improvement in the Raven colored progressive matrices in children with ADHD after WM training [34,35].

It should be noted that these studies on updating [45–47] did not include an active control group and thus did not control for non-specific effects of being in the training group (Box 1). Much greater transfer was observed in the study by Jaeggi *et al.* [47] than in the other two studies on updating [45,46]. This might be because of the dual-task paradigm or due to the adaptive paradigm used by Jaeggi *et al.*, whereby the amount of information to be kept in mind is gradually increased. The studies by Klingberg *et al.* [34,35] and Holmes *et al.* [37] identified training effects that were specific to the adaptive training compared to the non-adaptive training performed by the control groups. Load adaptation and extensive training (at least 3 weeks, or 8 h) could thus be two important factors for effective WM training.

Box 2. Methodological issues in neuroimaging of training

- Strategies can change during training. For example, a novice golfer might perform subvocal rehearsal of the instructions from his golf teacher while teeing up, but rely on non-verbal routines after a few months of practice. If the player were scanned during the course of practice, we would expect to see activation of language-related regions at the beginning of practice, followed by a gradual decrease in activation. This decrease in activity would probably be correlated with improved performance, but is not the underlying cause of the improved golfing performance, which presumably is to be found in the regions controlling motor functions of the arms and the body rather than the tongue. Similarly, performance of a novel task might initially require WM for the rules and stimulus-response mapping. Consistent stimulus-response mapping leads to automaticity and presumably decreased prefrontal activity.
- Repeated presentation of the same stimuli leads to priming, with faster identification, faster decisions and faster response time. Priming is consistently associated with decreased activity in sensory association areas, as well as the prefrontal cortex [80]. Although this priming is central in perceptual learning, it might be a parallel process not of primary interest in cognitive training studies.
- Changes in strategies, as well as priming, lead to a decrease in time-on-task. This is not the same as stating that metabolic brain activity per time unit decreases and that the brain is more 'efficient'. Relating brain activity to time-on-task is rendered more difficult by the fact the BOLD signal fluctuations detected by fMRI are non-linearly related to underlying brain activity [81,82].
- In defining 'practice', repeated performance, task-specific improvement and enhancement of an underlying ability are not the same. Generating a verb from a given noun might be faster the second time the nouns are heard (task-specific improvement) but this is not an indication that the vocabulary, fluency or verbal IQ (the underlying ability) has increased.

Neural correlates of WM training

Identifying the neural correlates of training-induced improvements has many caveats, since there are many parallel behavioral changes occurring during the course of training (Box 2). Furthermore, the aspects of brain activity associated with superior capacity are still a matter of debate. However, the majority of studies indicate a positive correlation between WM capacity and brain activity in task-relevant areas. Inter-individual differences in WM capacity have been positively correlated with activity in the intraparietal cortex [15–19] (but see also [48]). The WM capacity increase during childhood is also mainly positively correlated with brain activity in task-relevant areas of the intraparietal sulcus and prefrontal cortex [20–25] (but see also [49]). Conversely, the decline in WM during aging is mostly associated with a decrease in activity [50], although there might be different trends in different prefrontal areas [51]. In line with these results, neural network models suggest mechanisms by which higher BOLD activity would be associated with better capacity [26,27,52].

Table 2 summarizes functional magnetic resonance imaging (fMRI) studies that have investigated the effect of repeated performance of WM tasks on brain activity. Most of these studies only included short periods of repeated performance of WM tasks. Only four evaluated transfer to non-trained tasks, and are thus informative on the question of correlates of training-induced increase in capacity, and are summarized below.

The study by Olesen *et al.* [33] used the same method as in previous behavioral studies by Klingberg and collaborators [34,35]. In a first experiment, subjects were scanned multiple times before training and then once after training. Easy WM tasks were performed during scanning to achieve ceiling effects to avoid differences in behavior during scanning. In a subsequent experiment, subjects were scanned

Table 2. Neuroimaging studies of WM training[a]

Reference	Task	Amount of practice	Improvement in non-trained tasks	Findings
[57]	Face DMS	30 min	Not evaluated	↓ prec, parietal, occ
[73]	Object and spatial WM	40 min	Not evaluated	↓ prec, parietal, occ
[74]	Visuospatial DMS	45 min	Not evaluated	↓dlpfc, parietal, cing
[75]	Verbal WM	45 min	Not evaluated	↓ dlpfc, frontopolar
[76]	Object and spatial WM	2 h	Not evaluated	↓parietal, inf frontal
[77]	Visuospatial *n*-back	Asked to practice daily for 4 weeks	Not evaluated	↑parietal, dlpfc at 2 weeks ↓ after 4 weeks
[54]	Object WM	10 h over 10 days	For trained categories	↑ dlpfc, parietal ↓ cing, dlpfc, parietal
[53]	Updating tasks	11 h over 5 weeks	Updating tasks	↑ caudate ↓dlpfc, parietal
[33]	Visuospatial WM	12 h over 5 weeks	Verbal and visuospatial WM	↑ dlpfc, parietal, caudate ↓ cing
[55]	Verbal WM	20 h over 10 weeks	Verbal WM	↑ left inf frontal

[a]DMS, delayed matching to sample; prec, precentral sulcus; occ, occipital; dlpfc, dorsolateral prefrontal cortex; inf, inferior; cing, cingulate cortex; ↑, increase in activation; ↓, decrease in activation.

repeatedly during the course of training and the change in brain activity was correlated with the change in performance outside the scanner. Increased activity in prefrontal and parietal cortex was observed in both experiments (Figure 1a) and in the caudate nucleus in the second experiment.

In the study by Dahlin, the focus was on training of updating [53]. Analysis of the behavioral effect of this training revealed improvement in a non-trained updating task, but not in other WM tasks [45]. Training was associated with increased activity in the caudate nucleus (Figure 1b), but with decreased activity in parietal and prefrontal cortex. A group of older subjects lacked activity in the caudate nucleus at baseline, which was interpreted as the reason for the lack of transfer to the 3-back task in this group.

Moore and collaborators asked participants to practice categorization of complex visual objects [54]. Performance on a WM task with the trained categories of objects was compared to that for a WM task with novel objects. This training was thus focused on visual perception rather than WM capacity, but training resulted in higher accuracy for trained categories. Finally, the study by Wexler et al. included eight subjects with schizophrenia [55]. The three subjects who improved also showed increased prefrontal activity.

One consistent pattern in Table 2 is that short periods of training (<3 h) resulted in decreased brain activity,

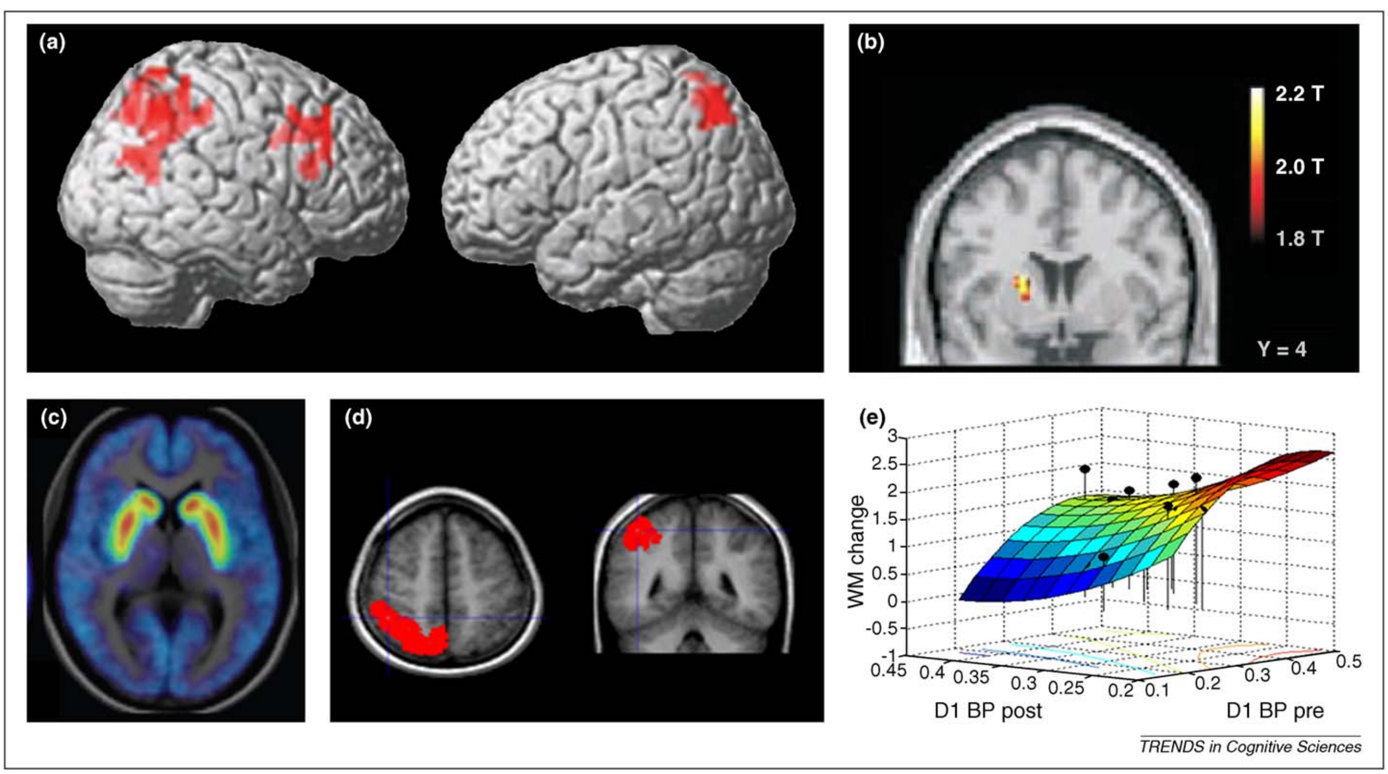

Figure 1. Training-related effects on brain activation and dopamine receptor density from neuroimaging studies. **(a)** Increases in frontal and parietal activity after training of WM (reproduced with permission from [33]). **(b)** Increased activity in the caudate nucleus after training of WM tasks requiring updating (reproduced with permission from [53]). **(c–e)** Results from study by McNab *et al.* [58]. **(c)** Density of dopamine D1 receptors. **(d)** Regions of interest based on activation during visuospatial WM tasks versus control tasks. **(e)** Relation between pre- and post-training measures of dopamine D1 receptors and gain in WM capacity based on the regions of interest specified in **(d)**.

whereas studies with longer periods of training showed a mixture of increases and decreases in different brain areas. One possible interpretation of this pattern is that increases in capacity are positively correlated with activity in the intraparietal cortex, middle and superior frontal gyri and caudate nucleus, but that this effect co-occurs with, and can sometimes be dominated by, decreases related to learning of strategies, priming during encoding and time-on-task effects (Box 2). A similar interpretation has been proposed for the effects of motor training, in which within-session effects were differentiated from long-term training effects [56]. The training effects in the study by Olesen *et al.* were localized to the prefrontal and parietal cortex, rather than to sensory association cortex [33]. These effects thus potentially indicate the involvement of the same multi-modal frontoparietal network. The frontoparietal effect might thus provide the basis for transfer between different WM tasks and to control of attention. Two studies found training-related changes in the basal ganglia [33,53], which have been associated with selection of relevant information in WM tasks [17] and might provide a more general mechanism that mediates improvement between different tasks. However, the methodological issues are complex and further studies are needed before any firm conclusions can be drawn. In particular, more studies that quantify the amount of priming and transfer and relate these directly to changes in brain activity are desirable. It would also be informative to separate cue- and response-related activity from maintenance activity [57].

A role for dopamine in cognitive training?

In addition to studying brain activity, McNab and colleagues investigated effects of WM training on the density of dopamine D1 and D2 receptors [58]. The dopamine system is of particular interest in WM training because dopamine is important for WM performance [59,60] and neuronal plasticity [61]. After 5 weeks of training, the increase in WM capacity for each subject compared to baseline was correlated significantly with changes in cortical D1 but not subcortical D2 receptors. This effect was mostly driven by a decrease in the number of D1 receptors, but was better explained by a non-linear, inverted-U function (Figure 1c–e). The decrease is consistent with animal studies showing that small amounts of D1-blocking agents can enhance WM-related activity, whereas large doses impair WM [62,63]. However, the causality in the study by McNab *et al.* is not clear. It is possible that intensive daily training on WM tasks resulted in increased endogenous release of dopamine during training, which led to an adaptive change in the density of dopamine receptors. However, it is also possible that dopamine plays a causal role and that WM training improves capacity partly by tuning dopaminergic transmission. If this is the case, transfer of training effects between cognitive functions might be explained not only by similarities in the underlying brain areas [33], but also by the underlying neurotransmitter systems they recruit, which is a different principle for translation of training effects to non-trained functions.

Support for the causal role of dopamine in inducing the training effect comes from a study indicating that a polymorphism in the DAT-1 receptor affects the outcome of training [64]. Unfortunately, the study was underpowered ($n = 29$) and the results did not reach significance, but might provide an interesting starting point for future studies of the genetics of cognitive plasticity. Knowledge of the biochemistry of training-induced plasticity could lead to new paradigms, combining cognitive training with pharmacological intervention.

Concluding remarks

WM training can induce improvements in performance in non-trained tasks that rely on WM and control of attention. This transfer effect is consistent with training-induced plasticity in an intraparietal–prefrontal network that is common for WM and control of attention. Adaptive training that focuses on control of attention could have similar effects and has shown promising results [65].

The observed training effects suggest that WM training could be used as a remediating intervention for individuals for whom low WM capacity is a limiting factor for academic performance or everyday life. The training-induced improvements observed in remembering an instruction or solving mathematical problems [37] underline the potential relevance of such training for education. However, training outside the laboratory setting involves many practical problems, such as assuring compliance over extended periods of training. For shorter training periods and without control of the training quality, there will be negligible effects, as illustrated in a study in which 10 min of unsupervised daily cognitive training three to four times per week did not result in any measurable cognitive effects [66].

There are many questions yet to be answered regarding WM training, such as the optimal duration and spacing of training to achieve transfer and durable improvements, the role of reward and motivation, age-dependent and other inter-individual differences in training potential, and upper limits for improvement.

The research on WM training reviewed in this article represents just the beginning of a new research field that explores the possibilities for enhancing cognitive functions by training. This field faces an even larger set of questions. Although plasticity is probably the rule, some functions might be more easily trainable than others. One study, for example, found that the training paradigm used for WM tasks did not work for training of inhibitory tasks [40]. Another challenge will be to explore individualized training regimes whereby training on several cognitive tasks is combined and to investigate combinations of cognitive training with pharmacological treatment and physical exercise.

Conflict of interest declaration

TK is a consultant for Cogmed Systems. This company provides software for working memory training used in several of the reviewed articles.

References

1 Conway, A.R. *et al.* (2003) Working memory capacity and its relation to general intelligence. *Trends Cogn. Sci.* 7, 547–552
2 Martinussen, R. *et al.* (2005) A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 377–384

3 Butterfield, E.C. *et al.* (1973) On the theory and practice of improving short-term memory. *Am. J. Ment. Defic.* 77, 654–669
4 Ericsson, K.A. *et al.* (1980) Acquisition of a memory skill. *Science* 208, 1181–1182
5 Miller, G.A. (1956) The magical number seven, plus-or-minus two or some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97
6 Buonomano, D.V. and Merzenich, M.M. (1998) Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186
7 Nudo, R.J. *et al.* (1996) Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J. Neurosci.* 16, 785–807
8 Recanzone, G.H. *et al.* (1992) Topographic reorganization of the hand representation in cortical area 3b of owl monkeys trained in a frequency-discrimination task. *J. Neurophysiol.* 67, 1031–1056
9 Rainer, G. and Miller, E.K. (2000) Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27, 179–189
10 Abikoff, H. and Gittelman, R. (1985) Hyperactive children treated with stimulants. Is cognitive training a useful adjunct? *Arch. Gen. Psychiatry* 42, 953–961
11 Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349
12 Linden, D.E. (2007) The working memory networks of the human brain. *Neuroscientist* 13, 257–267
13 Curtis, C.E. and D'Esposito, M. (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423
14 Klingberg, T. *et al.* (1996) Activation of multi-modal cortical areas underlies short-term memory. *Eur. J. Neurosci.* 8, 1965–1971
15 Vogel, E.K. and Machizawa, M.G. (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751
16 Todd, J.J. and Marois, R. (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754
17 McNab, F. and Klingberg, T. (2008) Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* 11, 103–107
18 Lee, K.H. *et al.* (2006) Neural correlates of superior intelligence: stronger recruitment of posterior parietal cortex. *Neuroimage* 29, 578–586
19 Gray, J.R. *et al.* (2003) Neural mechanisms of general fluid intelligence. *Nat. Neurosci.* 6, 316–322
20 Klingberg, T. *et al.* (2002) Increased brain activity in frontal and parietal cortex underlies the development of visuo-spatial working memory capacity during childhood. *J. Cogn. Neurosci.* 14, 1–10
21 Kwon, H. *et al.* (2002) Neural basis of protracted developmental changes in visuo-spatial working memory. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13336–13341
22 Olesen, P. *et al.* (2007) Brain activity related working memory and distraction in children and adults. *Cereb. Cortex* 17, 1047–1054
23 Scherf, K.S. *et al.* (2006) Brain basis of developmental change in visuospatial working memory. *J. Cogn. Neurosci.* 18, 1045–1058
24 Crone, E.A. *et al.* (2006) Neurocognitive development of the ability to manipulate information in working memory. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9315–9320
25 Ciesielski, K.T. *et al.* (2006) Developmental neural networks in children performing a categorical *n*-back task. *Neuroimage* 33, 980–990
26 Edin, F. *et al.* (2007) Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *J. Cogn. Neurosci.* 19, 750–760
27 Edin, F. *et al.* (2009) Mechanism for top-down control of working memory capacity. *Proc. Natl. Acad. Sci. U. S. A.* 106, 6802–6807
28 Baddeley, A. (2003) Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839
29 Engle, R.W. *et al.* (1999) Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In *Models of Working Memory* (Myake, A. and Shah, P., eds), pp. 102–134, Cambridge University Press
30 Corbetta, M. and Shulman, G.L. (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215
31 Vogel, E.K. *et al.* (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature* 438, 500–503
32 Kastner, S. *et al.* (2007) Topographic maps in human frontal cortex revealed in memory-guided saccade and spatial working-memory tasks. *J. Neurophysiol.* 97, 3494–3507
33 Olesen, P. *et al.* (2004) Increased prefrontal and parietal brain activity after training of working memory. *Nat. Neurosci.* 7, 75–79
34 Klingberg, T. *et al.* (2002) Training of working memory in children with ADHD. *J. Clin. Exp. Neuropsychol.* 24, 781–791
35 Klingberg, T. *et al.* (2005) Computerized training of working memory in children with ADHD – a randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 177–186
36 Kerns, K.A. *et al.* (1999) Investigation of a direct intervention for improving attention in young children with ADHD. *Dev. Psychol.* 16, 273–295
37 Holmes, J. *et al.* (2009) Adaptive training leads to sustained enhancement of poor working memory in children. *Dev. Sci.* 12, F9–15
38 Holmes, J. *et al.* (2009) Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Appl. Cogn. Psychol.* 12, 9–15
39 Westerberg, H. *et al.* (2007) Computerized working memory training – a method of cognitive rehabilitation after stroke. *Brain Inj.* 21, 21–29
40 Thorell, L.B. *et al.* (2009) Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113
41 Davidson, M.C. *et al.* (2006) Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078
42 Lezak, M.D. (1995) *Neuropsychological Assessment*, Oxford University Press
43 Korkman, M. *et al.* (1998) *NEPSY – A Developmental Neuropsychological Assessment*, Psychological Corporation
44 Mezzacappa, E. and Buckner, J.C. (2010) Working memory training for children with attention problems or hyperactivity: a school-based pilot study. *School Mental Health* DOI: 10.1007/s12310-010-9030-9
45 Dahlin, E. *et al.* (2008) Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychol. Aging* 23, 720–730
46 Li, S.C. *et al.* (2008) Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychol. Aging* 23, 731–742
47 Jaeggi, S.M. *et al.* (2008) Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6829–6833
48 Linden, D.E. *et al.* (2003) Cortical capacity constraints for visual working memory: dissociation of fMRI load effects in a fronto-parietal network. *Neuroimage* 20, 1518–1530
49 Schweinsburg, A.D. *et al.* (2005) fMRI reveals alteration of spatial working memory networks across adolescence. *J. Int. Neuropsychol. Soc.* 11, 631–644
50 Persson, J. and Nyberg, L. (2006) Altered brain activity in healthy seniors: what does it mean? *Prog. Brain Res.* 157, 45–56
51 Rajah, M.N. and D'Esposito, M. (2005) Region-specific changes in prefrontal function with age: a review of PET and fMRI studies on working and episodic memory. *Brain* 128, 1964–1983
52 Macoveanu, J. *et al.* (2006) A biophysical model of multiple-item working memory: a computational and neuroimaging study. *Neuroscience* 141, 1611–1618
53 Dahlin, E. *et al.* (2008) Transfer of learning after updating training mediated by the striatum. *Science* 320, 1510–1512
54 Moore, C.D. *et al.* (2006) Neural mechanisms of expert skills in visual working memory. *J. Neurosci.* 26, 11187–11196
55 Wexler, B.E. *et al.* (2000) Preliminary evidence of improved verbal working memory performance and normalization of task-related frontal lobe activation in schizophrenia following cognitive exercises. *Am. J. Psychiatry* 157, 1694–1697
56 Karni, A. *et al.* (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc. Natl. Acad. Sci. U. S. A.* 95, 861–868
57 Landau, S.M. *et al.* (2004) A functional MRI study of the influence of practice on component processes of working memory. *Neuroimage* 22, 211–221
58 McNab, F. *et al.* (2009) Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* 323, 800–802
59 Luciana, M. *et al.* (1992) Facilitation of working memory in humans by a D2 dopamine receptor agonist. *J. Cogn. Neurosci.* 4, 58–68
60 Muller, U. *et al.* (1998) D1- versus D2-receptor modulation of visuospatial working memory in humans. *J. Neurosci.* 18, 2720–2728

61 Stroemer, P.R.H. *et al.* (1998) Enhanced neocortical neural sprouting, synaptogenesis and behavioral recovery with D-amphetamine therapy after neocortical infarction in rats. *Stroke* 29, 2381–2395
62 Williams, G.V. and Goldman-Rakic, P.S. (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 376, 572–575
63 Vijayraghavan, S. *et al.* (2007) Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* 10, 376–384
64 Brehmer, Y. *et al.* (2009) Working memory plasticity modulated by dopamine transporter genotype. *Neurosci. Lett.* 467, 117–120
65 Rueda, M.R. *et al.* (2005) Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U. S. A.* 102, 14931–14936
66 Owen, A.M. *et al.* (2010) Putting brain training to the test. *Nature* DOI: 10.1038/nature09042
67 Wechsler, S. (1995) *WPPSI-RI. Wechsler Preschool and Primary Scale of Intelligence – Revised*, Psychological Corporation
68 Hossiep, M. *et al.* (1999) *Bochumer Matrizentest: BOMAT-Advanced-Short Version*, Hogrefe
69 Gronwall, D.M.A. (1977) Paced auditory serial-addition task: a measure of recovery from concussion. *Percept. Mot. Skills* 44, 367–373
70 Wechsler, D. (1981) *WAIS-R Manual*, The Psychological Corporation
71 Raven, J.C. *et al.* (1995) *Coloured Progressive Matrices*, Oxford Psychologist' Press
72 Wechsler, D. (1999) *Wechsler Abbreviated Scale of Intelligence (WASI)*, Harcourt Assessment
73 Landau, S.M. *et al.* (2007) Regional specificity and practice: dynamic changes in object and spatial working memory. *Brain Res.* 1180, 78–89
74 Garavan, H. *et al.* (2000) Practice-related functional activation changes in a working memory task. *Microsc. Res. Tech.* 51, 54–63
75 Jansma, J.M. *et al.* (2001) Functional anatomical correlates of controlled and automatic processing. *J. Cogn. Neurosci.* 13, 730–743
76 Sayala, S. *et al.* (2006) Increased neural efficiency with repeated performance of a working memory task is information-type dependent. *Cereb. Cortex* 16, 609–617
77 Hempel, A. *et al.* (2004) Plasticity of cortical activation related to working memory during training. *Am. J. Psychiatry* 161, 745–747
78 Oken, B.S. *et al.* (2008) Expectancy effect: impact of pill administration on cognitive performance in healthy seniors. *J. Clin. Exp. Neuropsychol.* 30, 7–17
79 Sternberg, R.J. (2008) Increasing fluid intelligence is possible after all. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6791–6792
80 Schacter, D.L. *et al.* (2007) Reductions in cortical activity during priming. *Curr. Opin. Neurobiol.* 17, 171–176
81 Buxton, R.B. *et al.* (1998) Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn. Reson. Med.* 39, 855–864
82 Zhang, N. *et al.* (2008) Investigating the source of BOLD nonlinearity in human visual cortex in response to paired visual stimuli. *Neuroimage* 43, 204–212