ORIGINAL ARTICLE

The mechanisms of far transfer from cognitive training: specifying the role of distraction suppression

Annie Desmarais[1,2] · François Vachon[1]

Abstract

Cognitive training aims to produce a durable transfer to untrained abilities (i.e., far transfer). However, designing effective programs is difficult, because far transfer mechanisms are not well understood. Greenwood and Parasuraman (Neuropsychol 30(6):742–755. https://doi.org/10.1037/neu0000235, 2016) proposed that the ability to ignore distractions is key in promoting far transfer. While the authors identified working-memory training based on the N-back task as an effective way to train distraction suppression, a recent meta-analysis concluded that this form of training rarely produces far transfer. Such inconsistency casts doubt onto the importance of distraction suppression in far transfer and calls for further examination of the role of this ability in cognitive training effectiveness. We propose here to conceptualize distraction suppression in the light of the load theory of attention, which distinguishes two mechanisms of distractor rejection depending on the level and type of information load involved: perceptual selection and cognitive control. From that standpoint, N-back training engages a single suppression mechanism, namely cognitive control, because it mainly involves low perceptual load. In the present study, we compared the efficacy of N-back training in producing far transfer to that of a new response-competition training paradigm that solicits both distraction suppression mechanisms. Response-competition training was the only one to produce far transfer effects relative to an active control training. These findings provided further support to Greenwood and Parasuraman's hypothesis and suggest that both selection perception and cognitive control need to be engaged during training to increase the ability to suppress distraction, hence to promote far transfer.

Introduction

Cognitive training represents an interesting opportunity to enhance cognitive abilities for people of every age. The aim of cognitive training is to transfer cognitive gains to untrained abilities, with the ultimate goal of improving a person's daily functioning. However, improvements were often limited to the trained task (learning effect). It was rare to see transfer to a different task measuring either the same cognitive process (near transfer) or another cognitive process (far transfer; Greenwood & Parasuraman, 2016; Simons et al., 2016). The lack of effectiveness of existing programs can be explained by the poor understanding of the mechanisms underlying far transfer (Greenwood & Parasuraman,

2016). Research on far transfer mechanisms became one of the top priorities in cognitive training domain, because it could allow the development of training programs that would promote long-term benefits of daily functioning for the population (Bherer, 2015; Foroughi et al., 2016).

Based on a review of the neuroscientific literature, Greenwood and Parasuraman (2016) proposed that the training of distraction suppression is central in producing transfer to fluid intelligence. Cowan (2001) affirmed that a good ability to suppress distractors facilitates the processing of working memory by protecting the relevant content from such irrelevant content. Greenwood and Parasuraman found a pattern of brain activation among several brain imaging studies that supports Cowan's theory. Indeed, analysis of neuroscientific and cognitive data suggested that suppression of distractors has the potential to generate benefits on working memory and fluid intelligence tasks. Specifically, Greenwood and Parasuraman argued that a training task that solicitates distraction suppression would alter the activation patterns of the dorsal attention network (intraparietal sulcus, superior parietal lobule and frontal eye field)

✉ Annie Desmarais
annie.desmarais.3@umontreal.ca

1 School of Psychology, Université Laval, Quebec City, Canada

2 Department of Psychology, Université de Montréal, 90 Avenue Vincent d'Indy, Montreal, QC H2V 2S9, Canada

and its variable relationship with default mode network. The authors discussed that mastering the training task can result in a dynamic reconfiguration of coherence in activity between specific brain networks involved in attention to external stimuli. These modifications promote transfer to other interconnected networks that involve additional cognitive processes, like fluid intelligence.

Greenwood and Parasuraman (2016) suggested that N-back training should be effective by soliciting distraction suppression and by triggering the activation of the dorsal attention network (Conway et al., 2001; Engle, 2002; Halford et al., 2007). However, N-back training has been the most studied in the literature (Linares et al., 2019), and there is no convincing evidence that such training can produce far transfer (Colom et al., 2013; Dougherty et al., 2016). Indeed, the multiple recent meta-analyses and literature reviews focusing on N-back training obtained a small or null effect size for the transfer measured on fluid intelligence, control of attention and other executive functions (Au et al., 2015; Brehmer et al., 2012; Karbach & Verhaeghen, 2014; Melby-Lervåg & Hulme, 2013; Melby-Lervåg et al., 2016; Morrison & Chein, 2011; Schwaighofer et al., 2015; Soveri et al., 2017; Von Bastian & Oberauer, 2014). Consistently, several researchers expressed their skepticism for the potential of N-back training to generate far transfer (Dahlin et al., 2008; Harrison et al., 2013; Linares et al., 2018, 2019; Redick et al., 2013) and some argued that training with the N-back task cannot ensure far transfer, according to the data reviewed (Colom et al., 2013; Dougherty et al., 2016). The inconsistencies between Greenwood and Parasuraman's hypothesis and empirical evidence on N-back training call into question the far transfer mechanism proposed by the authors and highlight the need to test empirically the role of distraction suppression in cognitive training effectiveness using another approach.

One type of training that had more success in producing far transfer is action video game training (AVG; Bavelier et al., 2012; Cohen et al., 2007). Murphy et al. (2016) used the load theory of attention (Lavie et al., 2004) to explain the benefits observed in AVG. They argued that perceptual capacity can be increased and that the load theory can be an interesting starting point for the development of an effective cognitive training. The load theory presents selective attention as arising from the involvement of two distinct distraction suppression mechanisms, which are engaged according to the level and type of information load involved in the task. When perceptual load is high, the perceptual capacity becomes saturated due to the processing of multiple stimuli. In such a context, a perceptual selection mechanism acts as a filter that excludes irrelevant stimuli based on physical distinctions such as location, colors, dimensions, or combination of features (Lavie et al., 2004). This is a rather passive mechanism since distractors are not treated simply because

there is not enough capacity to process them. When perceptual load is low, it is rather a cognitive control mechanism that actively rejects distractors (e.g., by inhibiting them), preventing them from undergoing more advanced processing to the detriment of priority information. Since it requires some mental effort, this form of selectivity is considered to be active.

In light of the load theory, the effectiveness of AVG-based training could lie in its complex environment, which can provide sometimes low, sometimes high perceptual load, thus engaging both distraction suppression mechanisms (Murphy et al., 2016). These authors proposed that this type of training could increase early filtering of irrelevant stimuli and control of attention. In fact, they claimed that far transfer could be favored by a training program promoting both low- and high-load contexts, but also encouraged further investigations to better understand this possible impact. We designed the present study in that perspective. In fact, we consider that far transfer to working memory tasks would be possible, due to the fact that both mechanisms are crucial for the protection of information in this cognitive process (Gaspar et al., 2016; Greenwood & Parasuraman, 2016; Vissers et al., 2017). The ineffectiveness of the N-back training, which involves presenting a single stimulus at a time, could be attributable to the fact that it offers low perceptual load only, hence stimulating only one of the distraction suppression mechanisms, namely cognitive control.

The present study

The goal of the present study was to determine whether the two putative mechanisms of distraction suppression put forward by the load theory play a crucial role in far transfer. To do so, we drew on the response competition (RC) task used to test the load theory to implement a cognitive training paradigm that allows training both perceptual selection and cognitive control and compare its effectiveness in inducing far transfer to that of the classical N-back training that engages only cognitive control.

The RC task developed by Lavie and Cox (1997) was used as a training task. This task consisted of presenting a circle formed of six letters with a distractor at the periphery or at the centre of the screen, as shown in Fig. 1. The circle was made up of a target letter, which the participants must find and identify, and five non-target letters. To stimulate both distraction suppression mechanisms, perceptual load was manipulated in the training task to provide a context of either high or low load. Perceptual load was controlled by changing the shape of the non-target letters in the circle. Low perceptual load display was composed of non-target letters that are homogeneous and dissimilar to the target letter in the circle of letters (the letters O with a target letter that
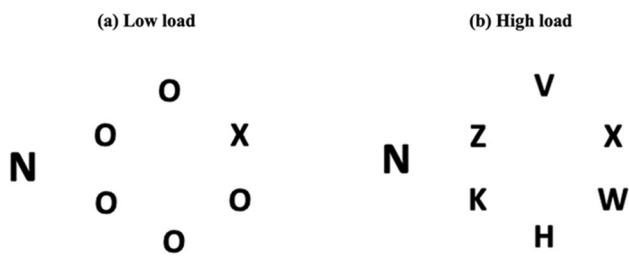
Fig. 1 Example of low-load and high-load displays in the response-competition task

is either X or N). According to Lavie et al. (2004), such a context does not saturate perceptual capacity, which favors perceptual processing of all letters presented on the screen, including distractors. These must then be actively inhibited by the cognitive control mechanism to prevent them from undergoing further processing to the detriment of the target letter. High load display was composed of various non-target letters similar to the target letter (the letters V, W, H, K, and Z with a target letter that is either X or N). This context prevented the processing of several stimuli, because the perceptual capacity quickly became saturated. Non-target letters were then discarded at an early stage in information processing by the perceptual selection mechanism that excluded irrelevant stimuli based on physical distinctions (Lavie et al., 2004). To our knowledge, such a cognitive training program based on the RC task has never been implemented in the scientific literature. Nevertheless, this was not the first cognitive training to adaptively manipulate distractor interference. For instance, Mishra et al., (2014, 2016) found far transfer as well as sensory plasticity following a training procedure whose mechanics specifically challenged participants to discriminate targets among progressively more interfering distractors. Such findings are promising as

they provide evidence that training approaches requiring the suppression of progressively more challenging distractors, such as the present RC training program, can successfully produce transfer effects.

The effectiveness of RC and N-bask trainings in producing far transfer was assessed relative to an active control training focusing on perceptual discrimination. Learning and transfer effects for each of the three training programs were evaluated by contrasting the same measures taken before (pretest) and after (posttest) the completion of the training sessions (see Fig. 2). These pretest–posttest experimental tasks were carefully selected to measure either selective attention or working memory. The two tasks measuring selective attention, the ability to process goal-relevant information while ignoring or inhibiting irrelevant information (Murphy et al., 2016), were the RC task and the auditory distraction task. The RC task involved searching for a target letter in the presence of visual distractors (e.g., Lavie et al., 2004), whereas the auditory distraction task required recalling the missing digit from a predefined series of digits in the presence of auditory distractors to ignore (e.g., Vachon et al., 2017). The two tasks assessing working memory, as the ability to "simultaneously store and manipulate information" (Greenwood & Parasuraman, 2016, p. 743), were the N-back task and the operation span (OSPAN) task. The N-back task consisted of recalling a target letter while inhibiting other target letters (e.g., Heinzel et al., 2014; Jaeggi et al., 2010), whereas the OSPAN task involved the serial recall of multiple target letters interspersed with a mathematical operation to be solved (e.g., Kane et al., 2007). As shown in Table 1, learning effects were evaluated using the task employed in the training phase, i.e. the RC task for RC training and the N-back task for N-back training. The auditory distraction task served as a measure of near transfer for RC training and far transfer for N-back training. Conversely, the OPSAN task

Fig. 2 Schematic representation of the pretest–posttest design employed in the present study. Participants were assigned to one of the three training programs: response competition (RC), N-back, or control. Blue shapes represent selective-attention tasks whereas pink shapes represent working-memory tasks
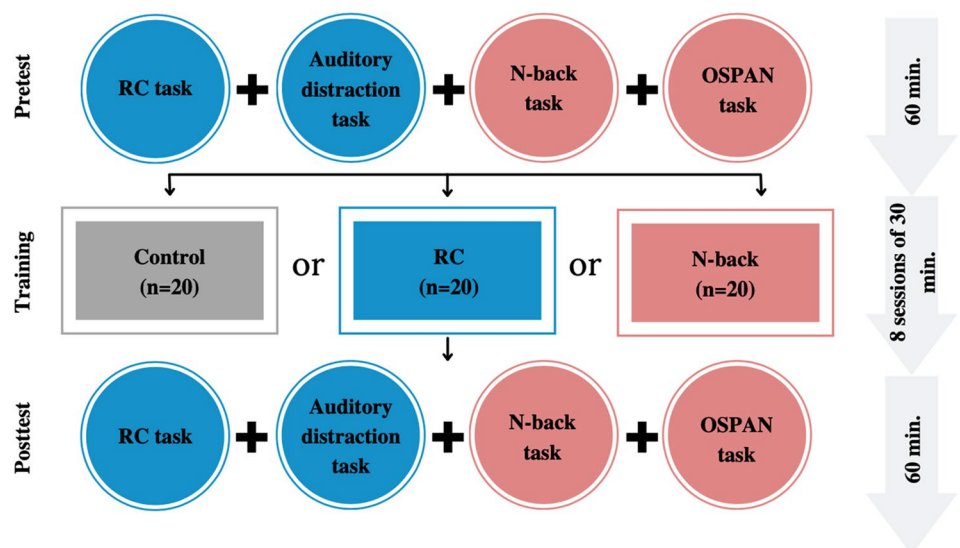
Table 1 Targeted tasks to measure learning, near transfer, and far transfer effects for response-competition (RC) and N-back trainings

Effect	RC training	N-back training
Learning	RC	N-back
Near transfer	Auditory distraction	OSPAN
Far transfer	N-back	RC
	OSPAN	Auditory distraction

Table 2 Sample characteristics ($n = 20$ for each group)

Training	Mean age Years (SD)	Gender (% female)
Control	24.10 (5.11)	55.00
RC	23.15 (4.82)	60.00
N-back	25.80 (5.66)	55.00

was used to evaluate near transfer for N-back training and far transfer for RC training.

Our hypothesis that training both perceptual selection and cognitive control should promote far transfer better than training a single distraction suppression mechanism predicts that RC training should produce greater far transfer effects than N-back training. More specifically, it is expected that RC training will induce significant posttest improvements in performance on the two working-memory tasks (i.e., N-back and OSPAN) relative to control training. Conversely, N-back training should produce little, if any, posttest improvements in performance on the two selective-attention tasks (i.e., RC and auditory distraction tasks) relative to control training, consistent with the existing literature (Colom et al., 2013; Dougherty et al., 2016).

Method

Participants

Seventy-three volunteers between the age of 18 and 40 years were recruited through emails and posters on the campus of Université Laval. Each volunteer reported normal or corrected-to-normal vision, normal hearing and no history of neurological disorders. The study was approved by the ethics committee of Université Laval and all participants provided written consent prior to the experiment. The financial compensation offered was $60 ($5 for each of the eight training sessions and $10 for each of the pretest and posttest sessions). Additionally, a contest participation coupon was offered for each completed session for a chance to win a $100 gift card.

Seven participants dropped out of the training program due to lack of time and six others were excluded due to data recording problems. Sixty were, therefore, included in the analysis of the present study. The final sample was composed of 90.40% undergraduate, graduate or postgraduate students at Université Laval. A total of 20 subjects were randomly assigned to each of the three trainings (Control, RC and N-back). As shown in Table 2, each of the training groups included in the analyses presented similar age and gender characteristics. Given a sample size of $n = 20$,

$\alpha = 0.05$, a training effect (i.e. a pretest–posttest difference) of size $\eta^2_p = 0.08$ could be detected with a statistical power of 0.80 according to G-Power software (Faul et al., 2007, 2009).

Design

The experimental design is illustrated in Fig. 2. The present study was conducted in laboratory and included a 60-min pretest phase followed by a training phase and a 60-min posttest phase. In both pretest and posttest, four experimental tasks were administered to assess selective attention (RC task and auditory distraction task; shown in pink in Fig. 2) and working memory (N-back task and OSPAN task; shown in blue in Fig. 2). During the training phase, participants performed one of the three training programs (N-back, RC, or Control). Training consisted in eight 30-min sessions that had to be completed within 4–8 weeks at a frequency of one or two sessions per week. Such flexibility in the training schedule was offered to promote study adherence and to avoid the loss of participants as much as possible due to time constraints.

Material

The presentation of instructions, stimuli and recording of responses for the N-back task, the N-back training, the RC task as well as the RC training were developed using Python. In addition, the presentation of instructions, stimuli and recording of responses for OSPAN task and auditory distraction task were programmed using E-Prime 2.0 software. For all tasks, visual stimuli were presented on a computer screen while auditory stimuli were presented through a headset. The experiment took place in an isolated dimly lit experimental room.

RC training

RC training was based on the RC task proposed in Lavie and Cox (1997) and was developed using different variations from different studies (Beck & Lavie, 2005; Forster & Lavie, 2008; Gupta & Srinivasan, 2010; Jenkins et al., 2005; Lavie & Cox, 1997; Lavie & De Fockert, 2003; Lavie et al., 2004;

Yeshurun & Marciano, 2013). The RC task consisted of a visual search task where participants must indicate which of the two target letters (N or X) is present by pressing a key (respectively, 0 or 2) on the keyboard as quickly as possible. The stimulus set was composed of a target letter and five non-target letters arranged in a circle at an equal distance from the centre of the screen (visual angle of 0.48° horizontally and 0.67° vertically at a distance of 60 cm; see Fig. 1). The target letter had an equal chance of being an X or N and appearing in any of the six positions of the circle. In addition, a distractor was positioned in one of the three following positions: on the periphery to the left or to the right of the circle of letters (visual angle of 0.57° horizontally and 0° vertically at a distance of 60 cm) or exactly at the centre of the circle. All stimuli (i.e., the fixation cross and the letters) were presented in black Times New Roman font on a white background.

Each trial began with a fixation cross lasting 600 ms, followed by the stimulus set which was presented for 100 ms or 150 ms. The screen then turned blank and participants had a maximum of 3000 ms to respond. When participants answered correctly, no feedback was given. When participants made a mistake or did not respond within 3000 ms, a pure tone was presented for 95 ms. The subsequent trial started automatically. A training session consisted of 12 practice trials followed by 15 blocks of 48 trials. An optional 5-min break was offered to participants after completing seven blocks to avoid cognitive fatigue.

Perceptual load varied within a block following the sequence HLLHHLLH, where "H" indicates a high-load trial and "L" indicates a low-load trial. In a counterbalanced order, half of the participants started with a high-load trial and the other half started with a low-load trial. In low perceptual load, the non-target letters were all Os. In high-load trials, the non-target letters corresponded to the letters H, K, V, W and Z, arranged randomly at the 5 remaining positions of the circle. The distractor could be congruent (either X for a target letter X and N for a target letter N), incongruent (N for a target letter X and X for a target letter N) and neutral (T for a target letter X or N, a face for a target letter X or N, and a cartoon character for a target letter X or N). Combinations of target letter identity, target letter position, distractor identity, and distractor position were equiprobable in each block of trials.

RC training had 18 difficulty levels created from variations of the RC task. As shown in Table 3, task difficulty was manipulated by varying the presentation time of the stimulus set (100 or 150 ms), as well as the perceptual load provided by the number of angular non-target letters in the high-load trials (K, H, V, Z, W), the nature (letter, face or cartoon) and position (periphery or fixation) of the distractor, and the presence of a mask following the stimulus set. The mask replaced each stimulus (target letter, non-target letters and distractor) by a "#", which reduced the visibility. The level of difficulty was adaptively increased depending on the performance of the participants within a given block. Each

Table 3 Characteristics of the difficulty levels of RC training

Level	Presentation time (ms)	Number of angular non-target letters in the high-load trials	Nature of the distractor for neutral trials	Position of the distractor	Presence of a mask following the stimulus set
1	150	1	Letter	Periphery	No
2	100	1	Letter	Periphery	No
3	150	3	Letter	Periphery	No
4	100	3	Letter	Periphery	No
5	150	5	Letter	Periphery	No
6	100	5	Letter	Periphery	No
7	150	5	Face	Periphery	No
8	100	5	Face	Periphery	No
9	150	5	Face	Fixation	No
10	100	5	Face	Fixation	No
11	150	5	Cartoon	Fixation	No
12	100	5	Cartoon	Fixation	No
13	150	5	Letter	Periphery	Yes
14	100	5	Letter	Periphery	Yes
15	150	5	Letter	Fixation	No
16	100	5	Letter	Fixation	No
17	150	5	Cartoon	Periphery	No
18	100	5	Cartoon	Periphery	No

training session began at the last level successfully achieved by the participants in the previous session. The threshold for moving to the next difficulty level was set according to a pilot study and existing literature. To reach the next level, participants must achieve fewer than 5% of errors and a block's average response time of less than 700 ms. When participants met these two criteria, they moved to a higher level. Otherwise, they stayed at the same level.

N-back training

N-back training was inspired from the training presented by Jaeggi et al. (2010) and Heinzel et al. (2014). Participants indicated if the letter presented at the center of the screen matches a letter N trials before, where N can take the value 2, 3, 4, 5, 6, 7, 8 or 9. All stimuli were presented in 120-point black Times New Roman font on a white background. The letters were selected from a set of 10 letters, namely F, H, K, L, M, Q, R, S, X, and Z (Green & Bavelier, 2003). A trial consisted in the presentation of a fixation cross for 1000 or 1800 ms, followed by a letter for 500 ms (see Fig. 3). Participants pressed the space bar to start a trial. When participants decided that a given letter matches the one presented in N previous positions, they had 500 ms to press the space bar. If they judged the letter was different, no action was required. When participants answered correctly, no feedback was given. A 95-ms pure tone was presented following an error. Each training session was composed of 25 test blocks of 20 + N letters. Participants were offered an optional 5-min break after completing 12 blocks to avoid cognitive fatigue.

Like RC training, the level of difficulty was adapted to the performance of the participants, to ensure maximum performance of the participants. N-back training consisted of 18 difficulty levels, as shown in Table 4. The value of N increased from 1 to 9 letters. Heinzel et al. (2014) established that the difficulty increases proportionally with

Table 4 Structure of N-back training (ISI = interstimulus interval)

Level	Task	ISI (ms)
1	$N = 1$	1000
2		1800
3	$N = 2$	1000
4		1800
5	$N = 3$	1000
6		1800
7	$N = 4$	1000
8		1800
9	$N = 5$	1000
10		1800
11	$N = 6$	1000
12		1800
13	$N = 7$	1000
14		1800
15	$N = 8$	1000
16		1800
17	$N = 9$	1000
18		1800

the value *N*, because the participants had to retain more letters in working memory. The highest level was set at $N = 9$ to avoid ceiling effects during training; the highest level achieved in cognitive training literature is usually $N = 6$ (Jaeggi et al., 2010). In addition, the value of the interstimulus interval (ISI) is 1000 ms for even levels and 1800 ms for odd levels. The difficulty also increased proportionally as the presentation time between stimuli increased, because the participants must retain the stimuli for a longer time (Heinzel et al., 2014). Each training session began at the last level successfully achieved by the participants in the previous session. To reach the next difficulty level, participants had to reach an accuracy rate of 95% without making any false alarms within a block.

Fig. 3 Illustration of the stimuli set shown in level $N = 2$ in the N-back task
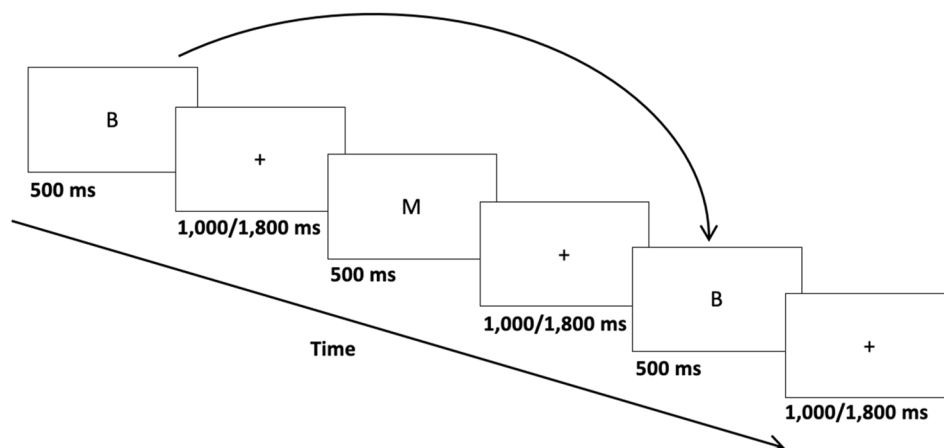
Active control training

The active control training program was based on a perceptual discrimination task taken from Lacouture (1997). More specifically, perceptual training was made up of three versions of the same relative judgment task (see Fig. 4): participants had to compare two lines on their length, width, or orientation. The lines were presented horizontally 10 cm from the center of the screen.

A trial consisted in the presentation of a black screen for 100 ms, followed by the stimulus set for 1500 ms, and then a white screen for 500 ms. Participants pressed the space bar to start a new trial. Participants had 2000 ms to indicate which line is the (a) longest, (b) thickest, or (c) closest to the vertical. Participants had to press the letter "C" when they decided the correct answer is the line on the left and the letter "M" when they considered the correct answer is the line on the right. When participants answered correctly, no feedback was given. A 95-ms pure tone was presented following an error.

A training session consisted of two practice trials and five consecutive blocks of 24 different trials for each of the three types of judgment. Participants underwent five blocks of one type of judgment at a time. The order of the three tasks was counterbalanced across participants and for the same participant across the eight sessions. An optional 5-min break was offered to participants after five completed blocks to avoid cognitive fatigue.

Each training session began at the last level successfully achieved by the participants in the previous session. The difficulty of each version of the task was manipulated by modifying the difference in length, width or angle from the vertical between the two lines. As shown in Table 5, the task has six levels of difficulty. Like N-back and RC trainings, the level of difficulty was adapted to the performance level of the participants. The first level begins with an 18% difference in length, width or angle from the vertical. Depending on the performance of the participants, the discriminability gradually decreases to 6%, which gradually increases the difficulty of the task (Matthews & Welch, 1997; Shiu & Pashler, 1992). The length, width or angle varies randomly with each type of judgment task, where the minimum and maximum value is 3.3 and 23.3 cm for the length, 1.5

Table 5 Structure of the active control training for the three judgment tasks

Level	Magnitude difference (%)	Maximal and minimal magnitude		
		Length (cm)	Thickness (cm)	Orientation
1	18	3.3–23.3	1.5–10.0	10.0°–170.0°
2	15			
3	13			
4	10			
5	8			
6	6			

and 10 cm for the width, and 10° and 170° for orientation. The greater magnitude is distributed randomly between the stimulus on the left and the one on the right. Participants reached a higher difficulty level in the block when they made two errors or less within a block.

Transfer tasks

Working memory tasks

At both pretest and posttest, working memory was measured using the n-back and OSPAN tasks.

N-back task This task was identical to the one used during N-back training, but it was restricted to levels $N = 2$ to $N = 6$. The higher level was set at $N = 6$ because it is usually the higher level reached in N-back training (Jaeggi et al., 2008). Participants completed two consecutive blocks of 24 trials for each of the five levels of the N-back task, which were presented from the lower level to the highest. The interstimulus interval was fixed at 1000 ms. The dependent variables were the percentage of correct responses and the response time for correct responses.

OSPAN task The task was to recall a series of letters in the order in which they were presented. Each letter was separated by a distractive task which consisted in solving a simple mathematical operation (see Fig. 5). All stimuli were shown at the center of the screen in black Times New Roman type on a white background. The OSPAN task was
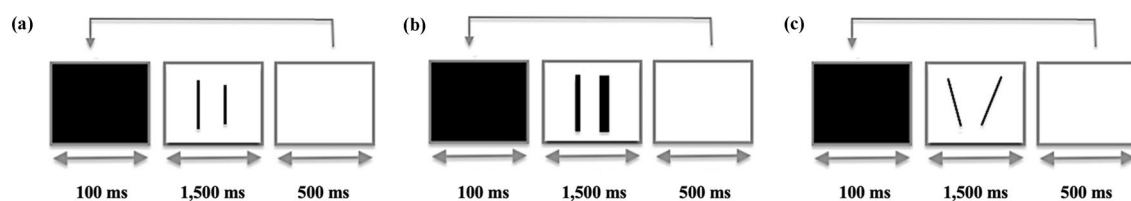
Fig. 4 Illustration of the stimulus set for the three relative judgment tasks of the control training: **a** length, **b** thickness, and **c** orientation

taken from the study of Foster et al. (2015) and was translated in French. We only modified the presentation time of the letters to 1000 ms.

Before performing the experimental phase, participants first familiarized with letter recall alone, mathematical problem solving alone (e.g., "$(1+0) \times 12 = ?$"), and then with the complete problem-letter sequence. After this familiarization phase, participants performed 15 experimental trials. In each trial, participants first solved a mathematical problem, and then saw a letter, and then solved another problem, and saw another letter, and so on. Such an alternation of mathematical problem and letter was repeated from 3 to 7 times for each trial with an unpredictable length each time. Each sequence length was encountered three times. At the end of the problem-letter sequence, participants had to recall, in order, the preceding letters by clicking on the corresponding letters presented in a 3×4 matrix (see Fig. 5). The dependent variable was the partial score (see Turner & Engle, 1989), computed by summing the number of letters correctly recalled in the correct order.

Selective attention tasks

At both pretest and posttest, selective attention was measured using the RC task and an auditory distraction task.

RC task Participants completed four versions of the RC task using Levels 6, 14, 16, and 18 from RC training. Participants completed one block of 48 trials for each version of the task, which were presented from the lowest level to the highest. Stimulus presentation was fixed at 100 ms. The dependent variables measured were the percentage of correct identification of the target letter and response time on correct trials. Because we had no a priori hypotheses about the potential influence of distractor compatibility on learning and transfer following RC training, we chose not to analyze distractor compatibility effects (e.g., Beck & Lavie, 2005) for the sake of brevity and clarity.
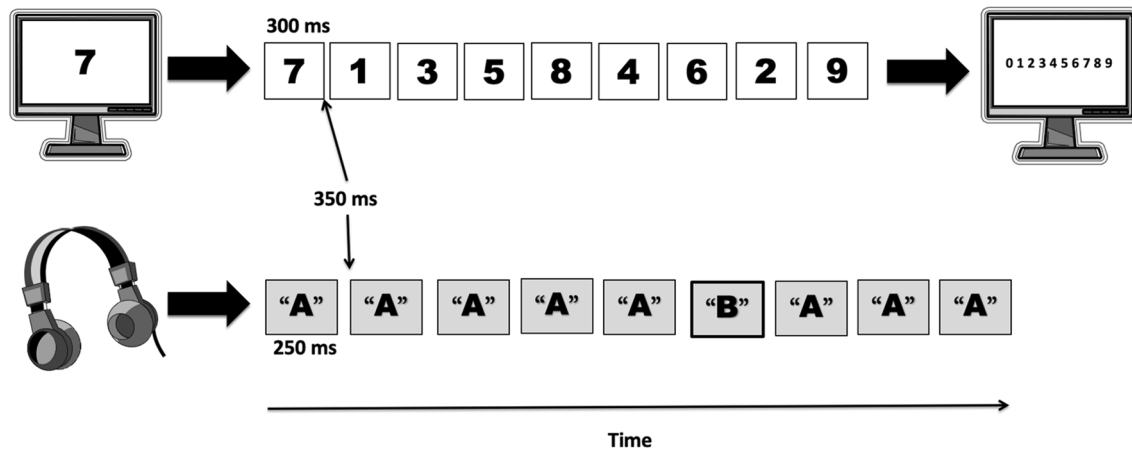
Auditory distraction task The auditory distraction task was inspired by the irrelevant sound paradigm of Vachon et al. (2017) and consisted of identifying the missing item from a list of visual stimuli while ignoring a sequence of auditory distractors presented simultaneously (see Fig. 6). Each visual sequence was composed of nine digits taken without replacement from the digit set 0–9 and arranged in a quasi-random order, with the constraint that successive digits were not adjacent integers. Each to-be-remembered item was approximately 2.39° in height and presented sequentially in black Times New Roman font at the centre of a white background for 300 ms with an interstimulus interval of 350 ms. The standard irrelevant auditory sequence consisted of nine repetitions of a spoken letter randomly selected from the set A–B–C–G–J–K–L–M–Q–S. On deviant trials, the sixth letter was replaced by one of the remaining letters, chosen at random (e.g., AAAAABAAA). Insertion of such a deviant sound tends to capture attention and confuse identification of the missing item (e.g., Vachon et al., 2017). The letters were recorded in a female voice at an approximately even pitch and edited using SoundForge (Sony Creative Software) to last 250 ms. The interstimulus interval was 350 ms.

On each trial, a to-be-remembered visual sequence and a to-be-ignored auditory sequence were presented concurrently. The auditory sequence began 125 ms before the beginning of the visual sequence. Following the disappearance of the last visual digit, the 10 digits appeared on the screen simultaneously in canonical order and the participants had to click on the digit that was not presented in the series (i.e., the missing item). Following two practice (control) trials, participants performed a single block of experimental trials containing 39 control trials and six deviant trials presented randomly with the constraint that deviant trials could not be adjacent. We computed a distractibility index as the dependent variable. It consists in the magnitude of the deviation effect, that is, the difference in the percentage of correct identification of the missing item between control and deviant trials. A positive score reflects the presence of distraction.
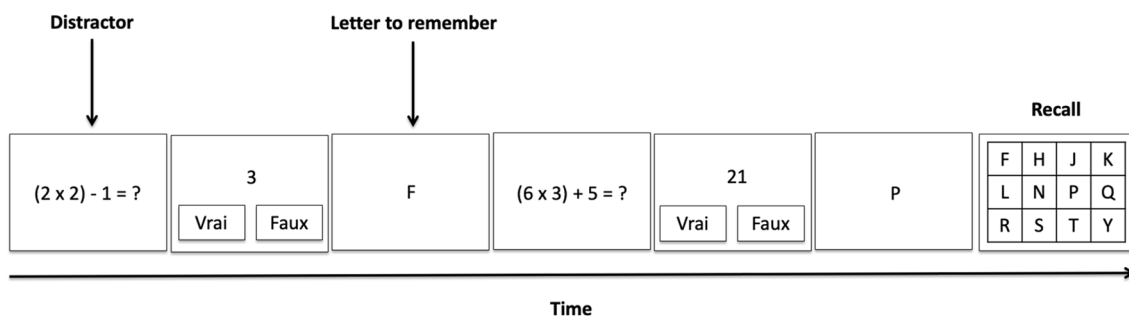
Fig. 5 Example of OSPAN. Here, two problem-letter sequences are illustrated

Fig. 6 Schematic presentation of stimulus presentation in the auditory distraction task. The top illustrates the to-be-remembered visual digits of the missing item task. The bottom depicts the to-be-ignored irrelevant sound sequence presented simultaneously through headphones. In this example, the auditory sequence contains a deviant item (here, a 'B' embedded among 'As'), indicated by the bold frame

Statistical analyses

Analyses were carried out on the average performance of the four transfer tasks at both pretest and posttest for each training program. Therefore, a 2×3 mixed ANOVA with Session (pretest vs. posttest) as the within-subject factor, and Training (control, RC and N-back) as the between-subjects factor was performed on each dependent variable. For each effect, we also reported the probability that the data favored the alternative hypothesis over the null hypothesis $[p_{BIC}(H_1|D)]$ as computed by a Bayes factor analysis using Masson's (2011) method. When the interaction was significant, simple effects tests were performed: dependent-sample t tests compared performance between pretest and posttest for each type of training using Holm–Bonferroni method. Data, analysis code, and stimuli are available from the corresponding author upon reasonable request.

Results

N-back task

Figure 7A illustrates the mean accuracy rate and Fig. 7B illustrates the average response time at the N-back task at pretest and posttest for each type of training. Accuracy seemed to improve at posttest following RC training and N-back training.

Accuracy

The ANOVA performed on accuracy rate revealed main effects of Session, $F(1, 57) = 25.72$, $p < 0.001$, $\eta^2_p = 0.311$, $p_{BIC}(H_1|D) > 0.999$, and Training, $F(2, 57) = 8.94$, $p < 0.001$, $\eta^2_p = 0.239$, $p_{BIC}(H_1|D) = 0.983$. In addition, there was a significant two-way interaction, $F(2, 57) = 20.28$, $p < 0.001$, $\eta^2_p = 0.416$, $p_{BIC}(H_1|D) > 0.999$. Simple effects tests revealed that participants in the N-back training group, $F(1, 19) = 52.70$, $p < 0.001$, $\eta^2_p = 0.735$, $p_{BIC}(H_1|D) > 0.999$, and in the RC training group, $F(1, 19) = 8.11$, $p = 0.010$, $\eta^2_p = 0.399$, $p_{BIC}(H_1|D) = 0.894$, improved significantly at posttest while no improvement was observed for the control group, $F(1, 19) = 1.29$, $p = 0.271$, $\eta^2_p = 0.063$, $p_{BIC}(H_1|D) = 0.299$.

Response time

The ANOVA performed on response time revealed that neither the main effect of Session, $F(1, 57) = 0.04$, $p = 0.849$, $\eta^2_p = 0.001$, $p_{BIC}(H_1|D) = 0.116$, nor that of Training, $F(2, 57) = 0.65$, $p = 0.525$, $\eta^2_p = 0.022$, $p_{BIC}(H_1|D) = 0.032$, was significant. The interaction was also not significant, $F(2, 57) = 2.08$, $p = 0.134$, $\eta^2_p = 0.068$, $p_{BIC}(H_1|D) = 0.121$.

OSPAN task

Figure 7C illustrates the mean partial score of the OSPAN task at pretest and posttest for each type of training. A visual inspection suggests that performance improved at posttest, but only following RC training. The ANOVA performed on these data revealed a significant main effect of Session, $F(1, 57) = 5.00$, $p = 0.029$, $\eta^2_p = 0.081$, $p_{BIC}(H_1|D) = 0.616$, but no main effect of Training, $F(2, 57) = 0.55$, $p = 0.580$, $\eta^2_p = 0.019$, $p_{BIC}(H_1|D) = 0.023$. The Session $\times$ Training interaction, $F(2, 57) = 2.21$, $p = 0.119$, $\eta^2_p = 0.072$, $p_{BIC}(H_1|D) = 0.135$, was not significant, which seems to contradict the visual impression that OSPAN performance did
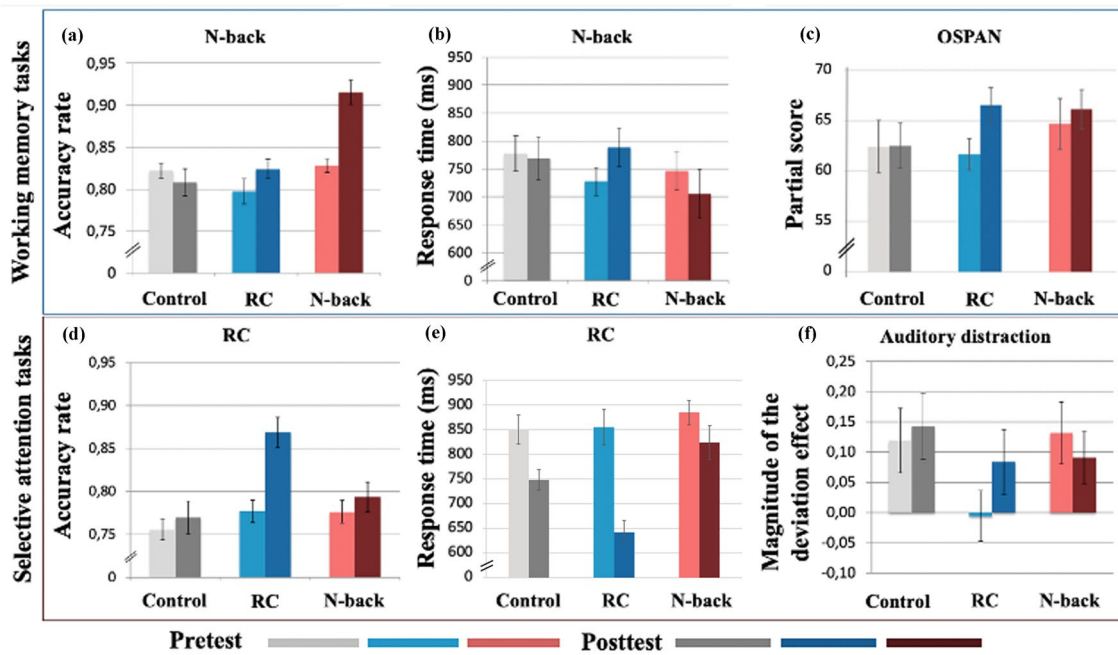
Fig. 7 Mean performance at pretest and posttest in each of the four transfer tasks (N-back, response competition [RC], OSPAN, and auditory distraction) for control (in grey), RC (in blue), and N-back (in red) training. Lighter colors represent pretest and darker colors represent posttest. Error bars represent the standard error of the means (color figure online)

improve at posttest in the RC training group (see Fig. 7C). To avoid making a Type II error, we did perform dependent-sample t tests to further investigate this issue. The tests revealed that partial score was indeed significantly higher at posttest following RC training group, $F(1, 19) = 13.27$, $p = 0.002$, $\eta^2_p = 0.411$, $p_{BIC}(H_1|D) = 0.978$. However, no pre-post difference was found for the N-back training group, $F(1, 19) = 0.57$, $p = 0.460$, $\eta^2_p = 0.029$, $p_{BIC}(H_1|D) = 0.231$, and the control group, $F(1, 19) < 0.01$, $p = 0.953$, $\eta^2_p < 0.001$, $p_{BIC}(H_1|D) = 0.183$. The fact that the Bayes factor for RC training provided strong evidence in favor of the alternative hypothesis suggests that there was a genuine improvement of OSPAN performance at posttest following RC training. The absence of significant interaction could be attributable to the relatively high variability found in the N-back and control groups. In fact, we plotted the pretest–posttest difference in partial score for each group (see Fig. 8) and noticed the presence of two extreme outliers (showing an increase of partial score over 20), one in control group and one in the N-back group. To reduce to potential influence of these outliers, we used the median of the pretest–posttest partial score difference instead of the mean. Whereas 15 participants for the RC training group were above the median, this was the case for only 5 and 8 participants in the control and N-back group, respectively. The median test contrasting the three groups was significant, $\chi^2(2) = 10.58$, $p = 0.005$, suggesting that the non-significant interaction likely arose from the presence of extreme outliers. Such results provided further evidence that
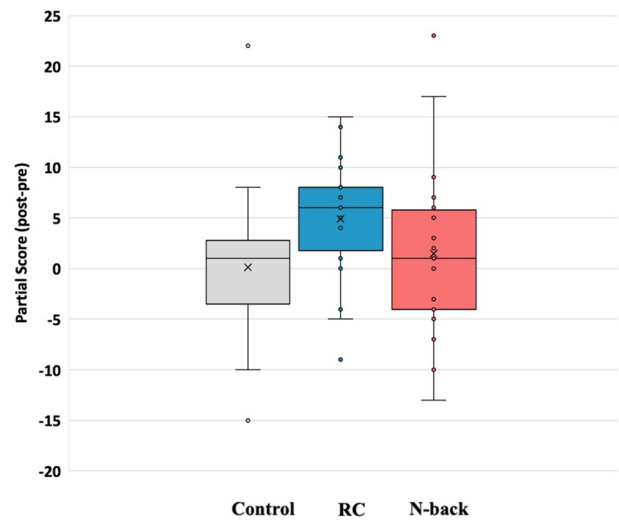
Fig. 8 Box plot of the pretest–posttest difference in OSPAN partial score for control, RC and N-back trainings

OSPAN performance improved at posttest, but only following RC training.

RC task

Performance on the RC task at pretest and posttest is plotted in Fig. 7D (mean accuracy) and Fig. 7E (mean response time) for each training. While response times seemed to

improve following each type of training, accuracy appeared to increase only after RC training.

Accuracy

The ANOVA performed on accuracy rate revealed significant main effects of Session, $F(1, 57) = 26.47$, $p < 0.001$, $\eta^2_p = 0.317$, $p_{BIC}(H_1|D) > 0.999$, and Training, $F(2, 57) = 4.82$, $p = 0.012$, $\eta^2_p = 0.145$, $p_{BIC}(H_1|D) = 0.641$. In addition, there was a significant interaction, $F(2, 57) = 10.10$, $p < 0.001$, $\eta^2_p = 0.262$, $p_{BIC}(H_1|D) = 0.994$. Simple effects tests indicated that the RC training group significantly improved at posttest, $F(1, 19) = 24.95$, $p < 0.001$, $\eta^2_p = 0.568$, $p_{BIC}(H_1|D) > 0.999$, whereas no such improvement was found for the N-back training group, $F(1, 19) = 2.66$, $p = 0.119$, $\eta^2_p = 0.123$, $p_{BIC}(H_1|D) = 0.445$, and the control training group, $F(1, 19) = 1.61$, $p = 0.219$, $\eta^2_p = 0.078$, $p_{BIC}(H_1|D) = 0.340$.

Response time

The ANOVA performed on mean response time revealed significant main effects of Session, $F(1, 57) = 88.86$, $p < 0.001$, $\eta^2_p = 0.609$, $p_{BIC}(H_1|D) > 0.999$, and Training, $F(2, 57) = 4.01$, $p = 0.023$, $\eta^2_p = 0.123$, $p_{BIC}(H_1|D) = 0.465$. In addition, there was a significant interaction between Session and Training, $F(2, 57) = 11.80$, $p < 0.001$, $\eta^2_p = 0.293$, $p_{BIC}(H_1|D) = 0.998$. Simple effects tests revealed that participants were significantly faster at posttest in each training group: control training $F(1, 19) = 38.44$, $p < 0.001$, $\eta^2_p = 0.669$, $p_{BIC}(H_1|D) > 0.999$, RC training, $F(1, 19) = 60.39$, $p < 0.001$, $\eta^2_p = 0.761$, $p_{BIC}(H_1|D) > 0.999$, and N-back training, $F(1, 19) = 6.48$, $p = 0.020$, $\eta^2_p = 0.254$, $p_{BIC}(H_1|D) = 0.808$. Given that all groups improved at posttest with regard to response time, we contrasted the pretest–posttest difference between each group to verify whether some groups showed better improvement than others. The tests revealed a significantly larger improvement following RC training relative to both control training, $F(1, 38) = 12.16$, $p = 0.001$, $\eta^2_p = 0.242$, $p_{BIC}(H_1|D) = 0.976$, and N-back training, $F(1, 38) = 17.07$, $p < 0.001$, $\eta^2_p = 0.318$, $p_{BIC}(H_1|D) = 0.997$. However, N-back and control training groups showed similar improvement in response time on the RC task, $F(1, 38) = 2.04$, $p = 0.161$, $\eta^2_p = 0.051$, $p_{BIC}(H_1|D) = 0.310$.

Auditory distraction task

Figure 7F shows the size of the deviation effect (i.e. the amount of distraction) at pretest and posttest in each training condition. A visual inspection of the plot suggests that distraction did not seem to diminish at posttest, regardless of the type of training. This was confirmed by the ANOVA carried out on these data, which revealed no main effect of Session, $F(1, 57) = 0.35$, $p = 0.557$, $\eta^2_p = 0.006$, $p_{BIC}(H_1|D) = 0.134$, nor of Training, $F(2, 57) = 1.91$, $p = 0.158$, $\eta^2_p = 0.063$, $p_{BIC}(H_1|D) = 0.104$. The interaction was also not significant, $F(2, 57) = 0.85$, $p = 0.431$, $\eta^2_p = 0.029$, $p_{BIC}(H_1|D) = 0.034$.

Discussion

The general objective of the study was to contribute to the understanding of far transfer mechanisms in cognitive training by empirically testing the hypothesis developed by Greenwood and Parasuraman (2016). They suggested that distraction suppression plays a central role in far transfer. The present study tested this hypothesis from the load theory perspective by examining whether a training program that solicits both cognitive control and perceptual selection—the two distraction suppression mechanisms proposed by Lavie et al. (2004)—would promote far transfer. We developed a novel training program using the RC task and compared its efficacy to that of a classical N-back training, which stimulates essentially only the cognitive control mechanism. The results showed that, relative to a control (perceptual) training, the only training program that produced far transfer was the one soliciting two distraction suppression mechanisms, namely RC training.

Far transfer mechanisms

We assumed that RC training involved both perceptual selection and cognitive control in suppressing distraction while N-back training solicited only the latter. The observation here that far transfer was restricted to RC training suggests that training both distraction suppression mechanisms plays a key role in far transfer. Given the newness of the RC training program and the fact that it was not closely match to the N-back training program, one must remain cautious, however, about the exact contribution of these two suppression mechanisms to far transfer effects. For instance, one cannot exclude the possibility that the success of RC training in producing far transfer relied exclusively on its cognitive control component.

Nevertheless, the present findings were consistent with the distraction suppression hypothesis which postulates that cognitive training programs designed to increase the ability to ignore distractors should promote far transfer. In light of existing studies, training both distraction suppression mechanisms would result in more effective working memory processes (Corbetta et al., 2008; Cowan et al., 2005). First, training the perceptual selection mechanism could allow for

a more effective suppression of distractors that appear concurrently to the target (see also Mishra et al., 2014, 2016). It would allow for a better selection of relevant information accessing working memory, thereby preventing the presence of irrelevant information that could interfere with it (Vissers et al., 2017). Moreover, various neuroimaging studies demonstrated the implication of perceptual selection in the efficacy of working memory (Gazzaley et al., 2005; Rees et al., 1997; Sreenivasan & Jha, 2007). For its part, training the cognitive control mechanism could allow for a more effective suppression of distractors that are actively retained in working memory (Corbetta et al., 2008). The effectiveness of such mechanism could be a determining factor for encoding and retention of information in working memory (Vissers et al., 2017). Moreover, Gaspar et al. (2016) showed that the ability to ignore distractors is linked to the capacity of working memory (see also Hughes et al., 2013; Sörqvist, 2010).

An unexpected result was the absence of near transfer from RC training to the auditory distraction task. Such an absence could be attributable to the inadequate calibration of the auditory distraction task. The results obtained indicated inconsistent and unexpected trends, limiting any possible conclusion for this task. First, contrary to what is usually found in this task (e.g., Hughes et al., 2013; Vachon et al., 2017), the deviation effect was absent from the pretest for the RC training group. This absence was surprising given that the deviation effect is considered a benchmark phenomenon for modelling short-term memory (Benchmark 6.3; Oberauer et al., 2018). It was therefore difficult to observe an improvement (i.e., a decrease in distraction) following the training program. In addition, the data appeared noisy, as suggested by the large error bars. This large variability in the data could be explained by the small number of deviant trials. There were only six deviant trials, which may not have been sufficient to provide a good signal-to-noise ratio. It is possible that a greater number of participants or trials would have been necessary to circumvent this limit. Hence, the parameters chosen in the auditory distraction task may have limited the possibility of observing a noticeable modulation of performance, calling into question the absence of near transfer. Another conceivable explanation for the lack of near transfer from RC training is the presence of a switch from visual to auditory distractors between the training task and the transfer task. In fact, there is evidence for the failure to observe transfer across sensory modalities (e.g., Schneiders et al., 2011, 2012, but see also Kattner et al., 2019). Even if this is a plausible hypothesis, it is important to mention that the literature on cross-modal transfer in cognitive training is inconsistent and further studies should be conducted to clarify this specific type of transfer (Taya et al., 2015).

Implications for Greenwood and Parasuraman's hypothesis

Greenwood and Parasuraman's (2016) hypothesis specified patterns of activation of brain regions, such as the activation of the dorsal attention network (see also Rueda et al., 2016), that characterize learning and far transfer. From this standpoint, the dorsal attention network, which involves the intraparietal sulcus, superior parietal lobule, and frontal eye fields, would be responsible for distraction suppression (see also Jerde et al., 2012). It is notably because N-back training is deemed to activate the dorsal network of attention that Greenwood and Parasuraman pointed to this form of training as a potentially effective way to induce far transfer. Yet the present study, consistent with previous research (De Simoni & Von Bastian, 2018; Melby-Lervåg & Hulme, 2013; Soveri et al., 2017), questioned this capacity to transfer N-back training to untrained abilities. In fact, the present findings suggest that an effective way to increase the ability to ignore distractions, hence to promote far transfer, is to train cognitive control (as in N-back training) along with perceptual selection. There is evidence that the stimulation of both distraction suppression mechanisms engages more cerebral regions than proposed by Greenwood and Parasuraman. Indeed, previous research found that the bilateral insula, left middle frontal gyrus and anterior cingulate cortex were activated during a task soliciting both cognitive control and perceptual selection (Torralbo et al., 2016; Wei et al., 2013). Moreover, both studies noticed that increasing perceptual load increase activity in the fronto-parietal regions (Torralbo et al., 2016; Wei et al., 2013). Wei et al. (2013) specified that this increased activation was located in the frontal eye fields, posterior parietal cortex, anterior insula and anterior cingulate cortex. Such expanded activation within the dorsal attention network pertaining to the solicitation of the perceptual selection mechanisms could be linked to the effectiveness of RC training relative to N-back training in transferring to untrained abilities observed in the present study. More research is required to corroborate this hypothesis.

Inefficacy of N-back training

The present study provided further evidence that N-back training fails to generate far transfer. The results were consistent with the existing literature showing that the benefits of N-back training are limited to the trained task (De Simoni & Von Bastian, 2018; Melby-Lervåg & Hulme, 2013; Soveri et al., 2017). It has been demonstrated that effective distraction suppression contributes to the effectiveness of working memory processes (Au et al., 2015; Karbach & Verhaeghen, 2014). We suggest that a training that solicit only the cognitive control mechanism would not be effective to produce

significant improvement on selective attention and working memory tasks, because such training would not increase distraction suppression at an early stage of information processing (Corbetta et al., 2008; Vissers et al., 2017). Indeed, it has been shown that effective distraction suppression at a perceptual stage can improve the processing of relevant stimuli in working memory (Gazzaley et al., 2005; Rees et al., 1997; Sreenivasan & Jha, 2007). If this hypothesis is accurate, then N-back training could be modified to engage both cognitive control and perceptual selection by integrating high and low perceptual load contexts. If such version of the N-back training proved to be effective in promoting far transfer, then one could conclude with more confidence that the lack of transfer from N-back training to other working memory and selective attention tasks observed in the present study was indeed due to the fact that it does not stimulate the perceptual selection mechanism.

Action video games

Well studied in the literature, AVG-based training is deemed to stimulate both distractor suppression mechanisms (Murphy et al., 2016). Indeed, AVG expose learners to a variety of tasks that alternate between a predominantly high perceptual load (e.g., attempting to shoot multiple opponents that appear on the screen at the same time) and predominantly a low perceptual load (e.g., attempting to shoot a single opponent without any distractor; Murphy et al., 2016). There is growing evidence that AVG-based training can produce far transfer by modifying the behavior and organization of neural networks (Altarelli et al., 2020). Nonetheless, it is not possible to establish what makes AVG effective in producing better performance at the perceptual, attentional, and memory levels (Green & Bavelier, 2003; Oei & Patterson, 2013). Based on the present study, we suggest that the solicitation of both distraction suppression mechanisms (cognitive control and perceptual selection) is at the basis of the effectiveness of AVG in enhancing untrained abilities (e.g., *Halo* or *Call of Duty*; Green & Bavelier, 2003, 2012). A parallel can be drawn between the findings from AVG studies and the results of the present study. Indeed, Mishra et al. (2011) provided neural evidence for enhanced suppression of distracting irrelevant information in AVG players relative to non-players. Furthermore, far transfer to selective attention and working memory tasks has been documented following AVG-based training (Cohen et al., 2007; Dye et al., 2009; Murphy et al., 2016). Specifically, researchers have shown that experts—who have played numerous hours of AVG—are more accurate and quicker at the RC task than beginners—who have played at best a few hours of AVG (Cohen et al., 2007; Dye et al., 2009). Moreover, experts are faster (Boot et al., 2008; McDermott et al., 2014) and more accurate at the N-back task compared to beginners (Colzato

et al., 2013). In fact, due to its success in producing transfer on selective attention and working memory tasks, AVG are being considered as an attractive therapy for the recovery of individuals with cognitive deficits in clinical neuropsychology (Nahum & Bavelier, 2020). Identifying the mechanisms underpinning the transfer capacities of AVG-based training to untrained abilities could help a better understanding of the outcomes from this new type of therapy.

Conclusion

The present research contributed to a better understanding of the mechanisms underlying far transfer in cognitive training, a priority in the cognitive enhancement field (Burgess et al., 2014; Greenwood & Parasuraman, 2016; Taya et al., 2015). The study provides support for the hypothesis that distraction suppression training fosters far transfer to working memory tasks in cognitive training. By borrowing the concept of cognitive control and perceptual selection to the load theory of attention (Lavie et al., 2004), the present study proposed that effective cognitive training—one that promotes far transfer—could benefit from engaging both distraction suppression mechanisms. The demonstrated capacity of RC training implemented in the present study to enhance performance to untrained tasks points to this novel training program as a promising tool that could be exploited by researchers in the field. Yet, more research is required not only to replicate the present findings, but also to investigate RC-training-related changes in neural activity.

Data availability Data and stimuli are available from the corresponding author upon reasonable request.

Code availability Analysis code is available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors have no known conflict of interest to disclose.

Ethics approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the University Ethics Committee.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent for publication Participants signed informed consent regarding publishing their data.

References

Altarelli, I., Green, C. S., & Bavelier, D. (2020). Action video games: From effects on cognition and the brain to potential educational applications. In S. Thomas, D. Mareschal, & I. Dumontheil (Eds.), *Educational Neuroscience* (1st ed., pp. 273–297). New York: Routledge. https://doi.org/10.4324/9781003016830

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review, 22*, 366–377. https://doi.org/10.3758/s13423-014-0699-x

Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. *Vision Research, 61*, 132–143. https://doi.org/10.1016/j.visres.2011.08.007.

Beck, D. M., & Lavie, N. (2005). Look here but ignore what you see: Effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 592–607. https://doi.org/10.1037/0096-1523.31.3.592

Bherer, L. (2015). Cognitive plasticity in older adults: Effects of cognitive training and physical exercise. *Annals of the New York Academy of Sciences, 1337*(1), 1–6. https://doi.org/10.1111/nyas.12682

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*(3), 387–398. https://doi.org/10.1016/j.actpsy.2008.09.005

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience, 6*, 63. https://doi.org/10.3389/fnhum.2012.00063

Burgess, G. C., Gray, J. R., Conway, A. R. A., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General, 140*(4), 674–692. https://doi.org/10.1037/a0024695

Cohen, J. E., Green, C. S., & Bavelier, D. (2007). Training visual attention with video games: Not all games are created equal. In H. O'Neil & R. Perez (Eds.), *Computer Games and Adult Learning* (pp. 205–227). Elsevier.

Colom, R., Román, F. J., Abad, F. J., Shih, P. C., Privado, J., Froufe, M., et al. (2013). Adaptive n-back training does not improve fluid intelligence at the construct level: Gains on individual tests suggest that training may enhance visuospatial processing. *Intelligence, 41*(5), 712–727. https://doi.org/10.1016/j.intell.2013.09.002

Colzato, L. S., van den Wildenberg, W. P., Zmigrod, S., & Hommel, B. (2013). Action video gaming and cognitive control: Playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychological Research Psychologische Forschung, 77*(2), 234–239. https://doi.org/10.1007/s00426-012-0415-2

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8*(2), 331–335. https://doi.org/10.3758/BF03196169

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron, 58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320*(5882), 1510–1512. https://doi.org/10.1126/science.1155466

De Simoni, C., & Von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. *Journal of Experimental Psychology: General, 147*(6), 829–858. https://doi.org/10.1037/xge0000453

Dougherty, M. R., Hamovitz, T., & Tidwell, J. W. (2016). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin & Review, 23*(1), 306–316. https://doi.org/10.3758/s13423-015-0865-9

Dye, M. W., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia, 47*(8–9), 1780–1789. https://doi.org/10.1016/j.neuropsychologia.2009.02.002

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Foroughi, C. K., Monfort, S. S., Paczynski, M., McKnight, P. E., & Greenwood, P. M. (2016). Placebo effects in cognitive training. *Proceedings of the National Academy of Sciences, 113*(27), 7470–7474. https://doi.org/10.1073/pnas.1601243113

Forster, S., & Lavie, N. (2008). Attentional capture by entirely irrelevant distractors. *Visual Cognition, 16*(2–3), 200–214. https://doi.org/10.1080/13506280701465049

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition, 43*(2), 226–236. https://doi.org/10.3758/s13421-014-0461-7

Gaspar, J. M., Christie, G. J., Prime, D. J., Jolicoeur, P., & McDonald, J. J. (2016). Inability to suppress salient distractors predicts low visual working memory capacity. *Proceedings of the National Academy of Sciences of the United States of America, 113*(13), 3693–3698. https://doi.org/10.1073/pnas.1523471113

Gazzaley, A., Cooney, J. W., Rissman, J., & D'esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience, 8*(10), 1298–1300. https://doi.org/10.1038/nn1543

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537. https://doi.org/10.1038/nature01647

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology, 22*(6), R197–R206. https://doi.org/10.1016/j.cub.2012.02.012

Greenwood, P. M., & Parasuraman, R. (2016). The mechanisms of far transfer from cognitive training: Review and hypothesis. *Neuropsychology, 30*(6), 742–755. https://doi.org/10.1037/neu0000235

Gupta, R., & Srinivasan, N. (2010). Distractor evaluation affects awareness under high load. *Journal of Vision, 10*(7), 197–197. https://doi.org/10.1167/10.7.197

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences, 11*(6), 236–242. https://doi.org/10.1016/j.tics.2007.04.001

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science, 24*(12), 2409–2419. https://doi.org/10.1177/0956797613492984

Heinzel, S., Schulte, S., Onken, J., Duong, Q. L., Riemer, T. G., Heinz, A., et al. (2014). Working memory training improvements and gains in non-trained cognitive tasks in young and older adults. *Aging, Neuropsychology, and Cognition, 21*(2), 146–173. https://doi.org/10.1080/13825585.2013.790338

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance, 39*(2), 539–553. https://doi.org/10.1037/a0029064

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning—Implications for training and transfer. *Intelligence, 38*(6), 625–635. https://doi.org/10.1016/j.intell.2010.09.001

Jenkins, R., Lavie, N., & Driver, J. (2005). Recognition memory for distractor faces depends on attentional load at exposure. *Psychonomic Bulletin & Review, 12*(2), 314–320. https://doi.org/10.3758/BF03196378

Jerde, T. A., Merriam, E. P., Riggall, A. C., Hedges, J. H., & Curtis, C. E. (2012). Prioritized maps of space in human frontoparietal cortex. *Journal of Neuroscience, 32*(48), 17382–17390. https://doi.org/10.1523/JNEUROSCI.3810-12.2012

Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 615. https://doi.org/10.1037/0278-7393.33.3.615

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological Science, 25*(11), 2027–2037. https://doi.org/10.1177/0956797614548725

Kattner, F., Samaan, L., & Schubert, T. (2019). Cross-modal transfer after auditory task-switching training. *Memory & Cognition, 47*(5), 1044–1061. https://doi.org/10.3758/s13421-019-00911-x

Lacouture, Y. (1997). Bow, range, and sequential effects in absolute identification: A response-time analysis. *Psychological Research Psychologische Forschung, 60*(3), 121–133. https://doi.org/10.1007/BF00419760

Lavie, N., & Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychological Science, 8*(5), 395–396. https://doi.org/10.1111/j.1467-9280.1997.tb00432.x

Lavie, N., & De Fockert, J. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception & Psychophysics, 65*(2), 202–212. https://doi.org/10.3758/BF03194795

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Linares, R., Borella, E., Lechuga, M. T., Carretti, B., & Pelegrina, S. (2018). Training working memory updating in young adults. *Psychological Research Psychologische Forschung, 82*(3), 535–548. https://doi.org/10.1007/s00426-017-0843-0

Linares, R., Borella, E., Teresa Lechuga, M., Carretti, B., & Pelegrina, S. (2019). Nearest transfer effects of working memory training: A comparison of two programs focused on working memory updating. *PLoS ONE, 14*(2), e0211321. https://doi.org/10.1371/journal.pone.0211321

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*(3), 679–690. https://doi.org/10.3758/s13428-010-0049-5

Matthews, N., & Welch, L. (1997). Velocity-dependent improvements in single-dot direction discrimination. *Perception & Psychophysics, 59*(1), 60–72. https://doi.org/10.3758/BF03206848

McDermott, A. F., Bavelier, D., & Green, C. S. (2014). Memory abilities in action video game players. *Computers in Human Behavior, 34*, 69–78. https://doi.org/10.1016/j.chb.2014.01.018

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270. https://doi.org/10.1037/a0028228

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer" evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*(4), 512–534. https://doi.org/10.1177/1745691616635612

Mishra, J., de Villers-Sidani, E., Merzenich, M., & Gazzaley, A. (2014). Adaptive training diminishes distractibility in aging across species. *Neuron, 84*(5), 1091–1103. https://doi.org/10.1016/j.neuron.2014.10.034

Mishra, J., Sagar, R., Joseph, A. A., Gazzaley, A., & Merzenich, M. M. (2016). Training sensory signal-to-noise resolution in children with ADHD in a global mental health setting. *Translational Psychiatry, 6*(4), e781. https://doi.org/10.1038/tp.2016.45

Mishra, J., Zinni, M., Bavelier, D., & Hillyard, S. A. (2011). Neural basis of superior performance of action videogame players in an attention-demanding task. *Journal of Neuroscience, 31*(3), 992–998. https://doi.org/10.1523/JNEUROSCI.4834-10.2011

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*(1), 46–60. https://doi.org/10.3758/s13423-010-0034-0

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin & Review, 23*(5), 1316–1340.

Nahum, M., & Bavelier, D. (2020). Video games as rich environments to foster brain plasticity. In N. F. Ramsey & J. D. R. Millán (Eds.),

Handbook of Clinical Neurology (pp. 117–136). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-63934-9.00010-X

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., et al. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin, 144*(9), 885. https://doi.org/10.1037/bul0000153

Oei, A. C., & Patterson, M. D. (2013). Enhancing cognition with video games: A multiple game training study. *PLoS ONE, 8*(3), e58546. https://doi.org/10.1371/journal.pone.0058546

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142*(2), 359. https://doi.org/10.1037/a0029082

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science, 278*(5343), 1616–1619. https://doi.org/10.1126/science.278.5343.1616

Rueda, M. R., Cómbita, L. M., & Pozuelos, J. P. (2016). Childhood and adolescence. In T. Strobach & J. Karbach (Eds.), *Cognitive Training* (pp. 33–44). Cham: Springer. https://doi.org/10.1007/978-3-319-42662-4_4

Schneiders, J. A., Opitz, B., Krick, C. M., & Mecklinger, A. (2011). Separating intra-modal and across-modal training effects in visual working memory: An fMRI investigation. *Cerebral Cortex, 21*(11), 2555–2564. https://doi.org/10.1093/cercor/bhr037

Schneiders, J., Opitz, B., Tang, H., Deng, Y., Xie, C., Li, H., & Mecklinger, A. (2012). The impact of auditory working memory training on the fronto-parietal working memory network. *Frontiers in Human Neuroscience, 6*, 173. https://doi.org/10.3389/fnhum.2012.00173

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*(2), 138–166. https://doi.org/10.1080/00461520.2015.1036274

Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics, 52*(5), 582–588. https://doi.org/10.3758/BF03206720

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17*(3), 103–186. https://doi.org/10.1177/1529100616661983

Sörqvist, P. (2010). High working memory capacity attenuates the deviation effect but not the changing-state effect: Further support for the duplex-mechanism account of auditory distraction. *Memory & Cognition, 38*(5), 651–658. https://doi.org/10.3758/MC.38.5.651

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*(4), 1077–1096. https://doi.org/10.3758/s13423-016-1217-0

Sreenivasan, K. K., & Jha, A. P. (2007). Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *Journal of Cognitive Neuroscience, 19*(1), 32–41. https://doi.org/10.1162/jocn.2007.19.1.32

Taya, F., Sun, Y., Babiloni, F., Thakor, N., & Bezerianos, A. (2015). Brain enhancement through cognitive training: A new insight from brain connectome. *Frontiers in Systems Neuroscience, 9*, 44. https://doi.org/10.3389/fnsys.2015.00044

Torralbo, A., Kelley, T. A., Rees, G., & Lavie, N. (2016). Attention induced neural response trade-off in retinotopic cortex under load. *Scientific Reports, 6*(1), 1–10. https://doi.org/10.1038/srep33041

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Vachon, F., Labonté, K., & Marsh, J. E. (2017). Attentional capture by deviant sounds: A noncontingent form of auditory distraction? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(4), 622–634. https://doi.org/10.1037/xlm0000330

Vissers, M. E., Gulbinaite, R., Van Den Bos, T., & Slagter, H. A. (2017). Protecting visual short-term memory during maintenance: Attentional modulation of target and distractor representations. *Scientific Reports, 7*(1), 1–14. https://doi.org/10.1038/s41598-017-03995-0

Von Bastian, C. C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research Psychologische Forschung, 78*(6), 803–820. https://doi.org/10.1007/s00426-013-0524-6

Wei, P., Szameitat, A. J., Müller, H. J., Schubert, T., & Zhou, X. (2013). The neural correlates of perceptual load induced attentional selection: An fMRI study. *Neuroscience, 250*, 372–380. https://doi.org/10.1016/j.neuroscience.2013.07.025

Yeshurun, Y., & Marciano, H. (2013). Degraded stimulus visibility and the effects of perceptual load on distractor interference. *Frontiers in Psychology, 4*, 289. https://doi.org/10.3389/fpsyg.2013.00289