Training enhances attentional expertise, but not attentional capacity: Evidence from content-specific training benefits

Roger W. Strong
Department of Psychology, Harvard University, Cambridge, MA, USA

George A. Alvarez
Department of Psychology, Harvard University, Cambridge, MA, USA

Cognitive training has become a billion-dollar industry with the promise that exercising a cognitive faculty (e.g., attention) on simple "brain games" will lead to improvements on any task relying on the same faculty. Although this logic seems sound, it assumes performance improves on training tasks because attention's capacity has been enhanced. Alternatively, training may result in attentional expertise—an enhancement of the ability to deploy attention to particular content—such that improvement on training tasks is specific to the features of the training context. The present study supported this attentional expertise hypothesis, showing that training benefits did not generalize fully from a trained attentional tracking task to untrained tracking tasks requiring a common attentional capacity, but differing in seemingly superficial features (i.e., retinotopic location and or motion type). This specificity suggests that attentional training benefits are linked to enhanced coordination between attentional processes and content-specific perceptual representations. Thus, these results indicate that shared attentional capacity between tasks is insufficient for producing generalized training benefits, and predict that generalization requires attentional expertise for content present in both training and outcome tasks.

Introduction

Cognitive training is a billion-dollar industry (Fernandez, 2013), fueled by consumers striving to maximize their cognitive abilities. By improving performance on training tasks, customers seek benefits in everyday life, such as heightened awareness while driving or better memory for coworkers' names. To produce such generalized benefits, practicing a training task must use and enhance mechanisms critical to untrained tasks (Ahissar & Hochstein, 2004). Therefore, training paradigms typically use training and outcome measures that putatively depend on common cognitive mechanisms (e.g., working memory tasks and fluid intelligence tasks; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008), hoping that training will improve performance on untrained tasks by enhancing a core cognitive ability, such as attention or memory. Although the logic of strengthening cognitive abilities in one context for use in another seems sound, examples of generalized training benefits are limited, and when found (Green & Bavelier, 2003; Jaeggi et al., 2008) often cannot be reproduced (Murphy & Spencer, 2009; Thompson et al., 2013). This inconsistency has generated criticism of training programs: In 2014 75 cognitive scientists released a statement declaring that little evidence exists supporting the efficacy of cognitive training products (Allaire et al., 2014), while in 2016 the cognitive training company Lumosity was fined $2 million for deceptive claims about the effectiveness of its products (Federal Trade Commission, 2016). This criticism highlights the need for a better understanding of the mechanisms enhanced by training.

In the present study, we trained participants on an attentional tracking task, then measured the degree of generalization to similar untrained tracking tasks that differed only in the features of the tracked items (translational vs. rotational movement, upper vs. lower visual field; see Figure 1). This paradigm allowed us to differentiate between two possibilities for how training might enhance attentional processing, which we label the "capacity-enhancing" and "expertise-building" hypotheses of attentional training. The capacity-enhancing hypothesis is the motivation behind cognitive training paradigms which posit that training increases a general attentional capacity's overall strength or effectiveness (e.g., Green & Bavelier, 2003). Analogous to how strength-training exercises (e.g., bench press)

Citation: Strong, R. W., & Alvarez, G. A. (2017). Training enhances attentional expertise, but not attentional capacity: Evidence from content-specific training benefits. *Journal of Vision*, *17*(4):4, 1–11, doi:10.1167/17.4.4.

doi: 10.1167/17.4.4 Received August 22, 2016; published April 7, 2017 ISSN 1534-7362

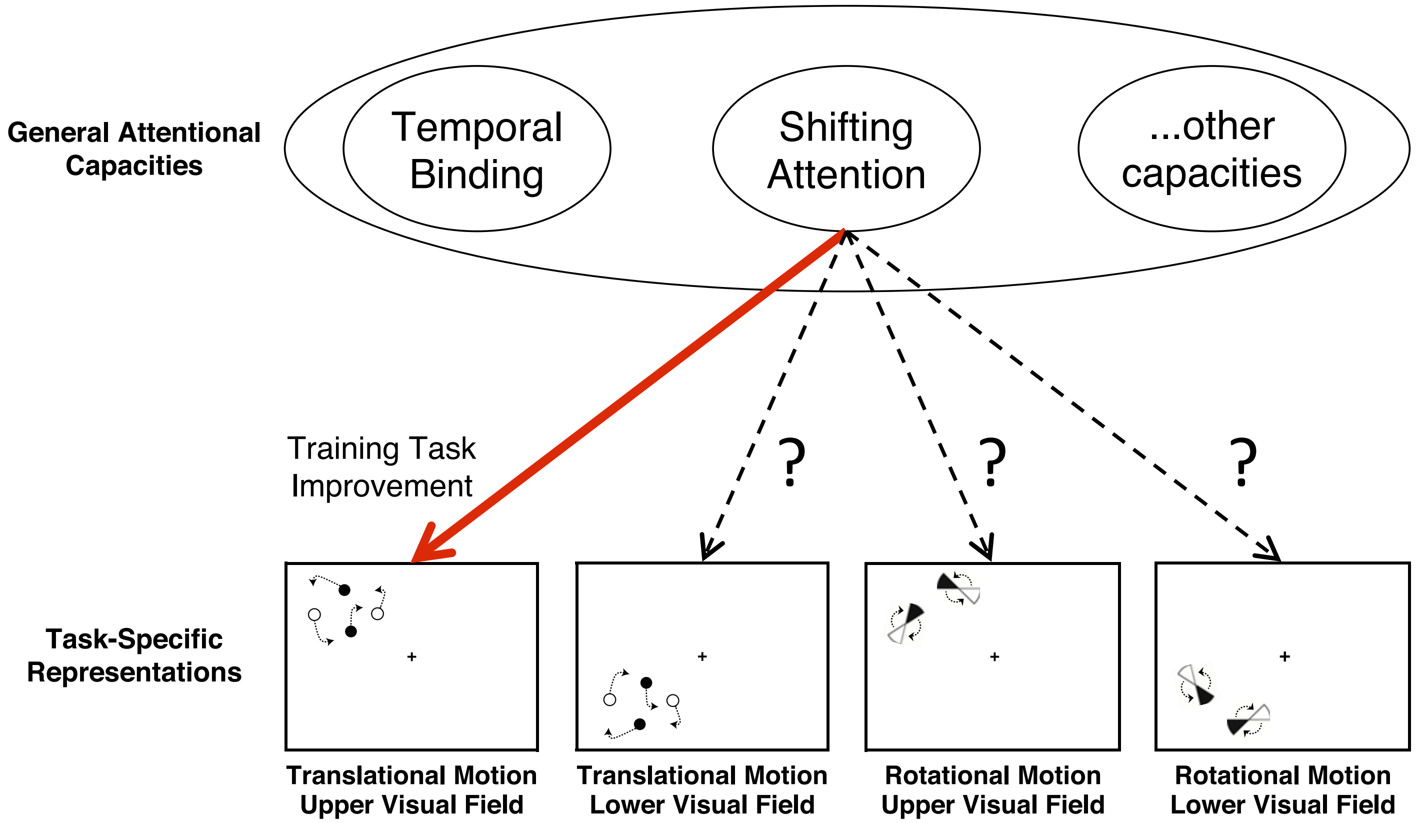

Figure 1. Example stimuli and possible training outcomes. Target items blinked white before fading to black as items began to move on screen (indicated by arrows). All of the tasks are known to share attentional capacities (based on the dual task experiment reported in Appendix A1). The top panel shows general attentional mechanisms, and highlights "shifting attention" as a possible general capacity shared by the tracking tasks. The arrows show that this general capacity is deployed to each of the tracking tasks (rotational, translational, upper visual field, lower visual field). The *capacity-enhancing hypothesis* assumes that training enhances general attentional capacities ("shifting attention" in this example), and therefore predicts that improvement on the training task (e.g., tracking translating dots in the upper visual field) should generalize to the untrained tasks, which are limited by a common attentional capacity. In contrast, the *expertise-building hypothesis* predicts that general mechanisms (e.g., "shifting attention") become more effectively deployed to representations specific to the training context (translational motion in the upper visual field in this example), and therefore predicts that improvement should be specific to the trained task (red arrow).

produce improved performance in untrained tasks that use the same muscles (e.g., lifting heavy furniture), the capacity-enhancing hypothesis predicts that improvement on training tasks should generalize to untrained tasks that use the same attentional capacity. In contrast, the expertise-building hypothesis posits that training does not improve general attentional capacities, but instead enhances the coordination between these capacities and task-specific representations (red arrow in Figure 1). In other words, the expertise-building hypothesis predicts that training allows an attentional capacity to make better use of training-specific representational content, without enhancing the capacity more generally. Analogous to how improving one skill requiring creativity (e.g., writing a poem) does not necessarily generalize to similar skills requiring creativity (e.g., writing a short story; Baer, 1996), the expertise-building hypothesis predicts that improvements on training tasks will fail to generalize to untrained tasks that do not require the attentional processing of content present during training, even when the trained and untrained tasks rely on the same attentional capacity (Figure 1).

We compared the predictions of the capacity-enhancing and expertise-building hypotheses by testing the transfer of training benefits between four very similar laboratory tasks (Figures 1 and 2). Each task is a variant of the multiple-object tracking paradigm (Cavanagh & Alvarez, 2005; Pylyshyn & Storm, 1988), which measures the ability to maintain attention simultaneously on multiple moving targets. Because multiple-object tracking is limited primarily by attentional processing (Alvarez & Franconeri, 2007; although a pre-attentive indexing of individual items may be required, Pylyshyn & Storm, 1988), it is a likely candidate to show generalization of training benefits under the capacity-enhancing hypothesis. This characteristic has led to the use of tracking tasks across a

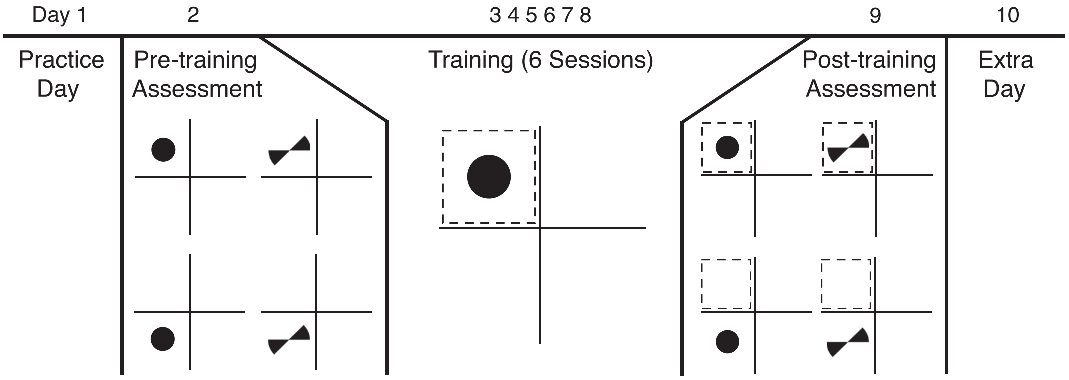

Figure 2. Training design. After pretraining assessment, participants trained for six days on one multiple-object tracking motion type in one location before completing a posttraining assessment (training on dot tracking in the upper left quadrant shown here, but counterbalanced across location and motion type). Control subjects completed same training design, except for taking days 3–8 off instead of training.

variety of cognitive training experiments. When used as an outcome measure for assessing the efficacy of other training paradigms, improvements in tracking performance have been found following video game training (Green & Bavelier, 2006), but were not found following improvement on a working memory task (Thompson et al., 2013). Following the completion of certain tracking training paradigms, benefits such as improved soccer performance (Romeas, Guldner, & Faubert, 2016), better biological motion perception (Legault & Faubert, 2012), and the enhancement of various cognitive functions (Parsons et al., 2014) have been reported, although other studies have failed to find generalization from tracking improvement to measures of working memory performance (Arend & Zimmer, 2012; Thompson et al., 2013). Additionally, tracking ability has been associated with performance in expert populations such as radar operators (Allen, McGeorge, Pearson, & Milne, 2004), professional athletes (Faubert, 2013), and laparoscopic surgeons (Harenberg et al., 2016), making it an intriguing task for comparing the predictions of the capacity-enhancing and expertise-building hypotheses of attentional training.

The four attentional tracking tasks used in the present study differed only in how the objects moved (translating dots vs. rotating pinwheels) and where the objects were located in the visual field (upper vs. lower visual field). Based on their similarity, it seems reasonable to assume that these tracking task share attentional resources. However, previous work has shown that two otherwise identical tracking tasks can draw on completely independent attentional resources when presented in separate hemifields (Alvarez & Cavanagh, 2005). Thus, in order to ensure that these tasks are attentionally demanding and share common attentional resources, we ran a preliminary experiment using a dual-task method (the attentional operating characteristic, Sperling & Melchner, 1978; see Appendix A1). The results of this experiment showed that completing any pair of the tracking tasks simultaneously resulted in a direct performance tradeoff between the tasks (i.e., participants had to perform worse on one task to do better on the second), unlike tasks that draw from independent capacities (tracking in the left vs. right hemifield; Alvarez & Cavanagh, 2005) or partially independent capacities (tracking vs. search; Alvarez, Horowitz, Arsenio, DiMase, & Wolfe, 2005). Thus, these preliminary results show that all the tracking tasks used in our training study are limited by a common attentional capacity (or possibly multiple common capacities critical to tracking, such as the capacities for shifting attention and sustaining attention). Therefore, any improvement in the trained task that is the result of an enhanced attentional capacity would generalize to the other tasks.

Having established that the four attentional tracking tasks in the present study are limited by a common attentional capacity (or multiple common capacities), we trained participants using either the translating dots or the rotating pinwheels motion type, and compared their performance to a no-training control group to account for practice effects. If training enhances attentional capacity (capacity-enhancing hypothesis), we expect complete generalization between the four tracking tasks, as each shares an attentional capacity critical to tracking performance. However, although we verified that these four tasks are limited by a common attentional capacity, each uses different representational mechanisms specific to translational versus rotational motion (Morrone et al., 2000) and to the representation of the upper versus lower visual field (Holmes & Lister, 1916; Sereno et al., 1995). Thus, if training enhances attention's coordination with content-specific representations (expertise-building hypothesis), we expect improvement to be specific to the training context (i.e., to the location and type of motion practiced during training).

Method

Procedure

Over a 10-day training design (Figure 2), participants were required to track moving objects with their attention while keeping their eyes focused on a central fixation point (Figure 1). Following an initial practice day, pretraining speed thresholds were determined for four different tracking tasks (translating dots and rotating pinwheels, in the upper and lower visual fields). Participants next completed six days of training, practicing one motion type in one retinotopic location (e.g., translating dots in the top left quadrant). Following the training period, posttraining speed thresholds were determined for the four tracking tasks. This design allowed analysis of whether performance benefits for the trained motion type in the trained retinotopic location (trained condition) generalized to the untrained motion type in the trained location (new motion condition), the trained motion type in the untrained location (new location condition), or the untrained motion type in the untrained location (both new condition).

On day one of the experiment, participants completed approximately 30 minutes of practice trials, containing a mix of the two motion types in either the upper-left and lower-left or the upper-right and lower-right visual quadrants. Participants were required to take a short break 15 minutes into this practice session. On day two of the experiment, pretraining tracking speed thresholds were obtained for both motion types (translating dots and rotating pinwheels) in two retinotopic locations (the upper and lower visual field on the opposite left/right side of the screen from the practice day). Two randomly interleaved staircases were completed (10 reversals—see Appendix A2 for full description of staircase parameters) for each of the four multiple-object tracking motion type/location combinations (resulting in eight total staircases being completed). For each staircase, we estimated the speed threshold by averaging the final four reversal speeds, resulting in two speed-threshold estimates per motion type/location combination; the mean of these two speed-thresholds was used as the baseline speed for each motion type/location combination. Participants were required to take a short break every 20 minutes until all staircases were completed.

For days 3–8, participants completed approximately 55 minutes of trials of one motion type in one location (e.g., dots in the upper left of the screen). The trained motion type and location for each participant were selected prior to beginning the experiment, and were counterbalanced across the two motion types and four visual quadrants. Participants were required to take a short break after 20 minutes, and again after 40 minutes. Each training day, tracking speeds for each trial were calculated using two randomly interleaved staircases (see Appendix A2 for details). At the conclusion of each training session, the mean of the two staircases' final four reversals were used to calculate the initial trial speeds for the following training day.

After six days of training, posttraining speed thresholds were determined using the same procedure as the pretraining assessment. The starting staircase speeds were the same for the posttraining assessment (day 9) as they were for the pre-training assessment (day 2). In pilot experiments, we found that observers often showed a drop in performance on what they believed to be the last day of testing. Thus, participants completed an additional one-hour session (day 10) after the posttraining assessment, so that the critical post-training measure (day 9) would not occur on the final day of the experiment. Because our pilot data suggested that these "extra day" data should not be analyzed, we only tested participants for one-hour on day 10, which was not enough time for all of the staircases to asymptote (i.e., not enough reversals could be completed in an hour to provide comparable speed-threshold estimates).

Control participants completed the same practice and assessment sessions as the training participants, but completed no training between the pretraining and posttraining assessment; these control participants did not come into the lab or have any contact with the lab during the interval between the two assessments. Although no-contact control groups are a problem for cognitive training studies claiming generalization of training benefits (Boot, Simons, Stothart, & Stutts, 2013), such a strategy is actually advantageous when reporting specificity of training benefits, as it allows measurement of the amount of the post-training improvement attributable to taking the assessment a second time. The source of any improvement following an active-control task would be less clear, and could potentially be the result of enhanced attentional capacity induced by the control task, complicating claims about the specificity of training benefits. Although we cannot rule out the possibility that the no-contact control participants displayed greater motivation than the training participants following their extended break (potentially producing performance gains beyond test-retest improvement), this possibility is very unlikely given that active controls typically perform more similarly to training subjects than no-contact controls (Melby-Lervåg & Hulme, 2013; Au et al., 2015).

Stimuli

Two different multiple-object tracking stimulus types were used during the experiment (Figure 1), both of which required tracking two target items moving among two distractor items for 6 seconds (510 frames at 85 Hz; 40 cm × 30 cm CRT monitor; 50 cm viewing distance, 122 cd/m^2 background luminance). One task required tracking two of four moving dots, while the other required tracking one spoke on each of two rotating pinwheels (each pinwheel had two spokes, one of which had to be tracked).

Each trial occurred in one quadrant of the screen. The black translating dots (diameter $= 1.1°$) were contained within a $14.8° \times 14.8°$ invisible box, centered at 45° between the horizontal and vertical midline along an invisible circle of radius 13.4°. A minimum spacing of 2.2° was maintained between the dots. The black pinwheels (width $= 5.1°$) were contained within the same $14.8° \times 14.8°$ invisible box, with each pinwheel centered 22.5° from the center of the box along the invisible circle of radius 13.4°. The pinwheels rotated and changed direction independently of one another, with changes in spin direction occurring a minimum of 75 frames apart and a maximum of 400 frames apart; each pinwheel had a 1/200 chance of changing spin direction at each frame between the minimum and maximum. At the start of each trial, two of the four items were designated as targets by blinking. The target items then gradually faded to the color of the distractor items during the first second of motion. During the 6 seconds of motion, subjects were required to maintain central fixation by looking at a crosshair (width $= 0.6°$) at the center of the screen. If participants broke central fixation ($>2°$ from center, monitored using EyeLink 1000, SR Research, Ottawa, ON, Canada), the trial was terminated and not used for calculating the participant's tracking-speed threshold. Blinking during the 6 seconds of motion did not count as a broken fixation. After 6 seconds, the targets stopped moving, and participants were required to try to identify the two target items by selecting them with a mouse click. Trials were only marked as correct when both target items were successfully identified.

Due to experimenter error, JPEG compression of the pinwheel stimuli caused the target items to appear slightly darker (7.2 cd/m^2) than the distractor items (8.5 cd/m^2) for the pinwheel stimuli for the dot training group. This error did not occur for the black (5.8 cd/m^2) dot stimuli, and was not present for pinwheel training group or the control group. Although a potential concern, we are confident that this error did not influence the results of the dot training group (the only group exposed to the error) for several reasons. First, the difference was subtle, and was reported by only one observer who completed an experimental session with near perfect accuracy, despite the pinwheels spinning at impossibly fast speeds ($>1200°$/s; this subject was removed from the experiment). No other observers appeared to notice this error, as none achieved such high performance in any experimental session. Second, if participants had noticed this error and performed well on the pretraining task, then it would artificially look like they improved less on the posttraining tasks. However, the stimulus error did not appear to result in inflated performance in the pretraining assessment for the pinwheel stimuli: An independent-samples t test revealed that the dot training group's baseline performance for pinwheels with the error ($M = 291.3°/s$, $SD = 85.3°/s$) and the pinwheel training group's baseline performance for pinwheels without the error ($M = 292.4°/s$, $SD = 74.9°/s$) was equivalent, $t(30) = 0.04$, $p = 0.97$. Finally, although this error could have artificially produced increases in the pinwheel transfer conditions (new motion and both new) relative to the dot conditions (trained and new location) if the dot training group had picked up on the error in the posttraining assessment, the dot training grouped displayed greater gains for the dot-tracking conditions than for the pinwheel-tracking conditions (meaning this outcome would have worked against our findings of specificity). In short, with the exception of one subject, participants did not appear to notice this stimulus error, and if they had it would work against the ultimate conclusions drawn based on data from both training groups combined.

Participants

Participants ($N = 48$, $M_{\text{age}} = 23.5$; 21 female, 27 male) were recruited from Harvard University and the Cambridge, Massachusetts community via posters describing the study. Participation was limited to individuals between the ages of 18–35 with normal or corrected-to-normal vision; those with corrected-to-normal vision were required to wear contacts during their participation. Participants received \$10/hour compensation, plus an additional \$25 completion bonus at the end of day 5 (for dot training and pinwheel training subjects) and day 10. To increase participant engagement, participants also received performance-based bonus pay of up to \$5 per day.

Before beginning subject recruitment, a target sample of 16 participants for each of the three groups (dot training, pinwheel training, control) was decided upon for inclusion in our final analysis. This number was selected after examining effect sizes of interest from pilot data. Also before beginning recruitment, we decided to replace any subjects in the dot training and pinwheel training groups who failed to display at least 25% improvement in at least one of the four experimental conditions (the criterion was applied to all conditions so that we would not bias the sample toward participants who showed spuriously high improvement in the training condition). Because the study's goal was to examine whether training benefits would generalize, we wanted to include only participants who showed at least moderate training benefits. Control participants, however, were not replaced for failing to improve by 25% in at least one of the four tasks; although this may have slightly underestimated the amount of test–retest improvement in the control condition relative to the four training conditions, this was a conservative analysis strategy given our conclusions of training specificity. Two subjects (one dot training and one pinwheel training subject) were replaced for failing to improve by 25% in at least one experimental condition. An additional five subjects (two dot training, one pinwheel training, and two control subjects) were replaced for having an excessive amount of fixation breaks during their baseline or final assessment session (>3 standard deviations from the mean broken fixations during assessment), indicating extreme difficulty keeping fixation. Finally, a one dot training subject was excluded after being found using a cell phone during trials of the final assessment. Seven additional subjects were replaced after dropping out of the experiment during the training period (three dot training, one pinwheel training, and three control participants). Had subject replacement not occurred, the only difference in significance testing would have been between the new location condition and the control condition, which would not have been significantly different without subject replacement (see Appendix A3). Although this lack of a difference between the new location and control conditions would have actually better supported our claims of training specificity, the exclusion criterion was determined before running the study, and we believe our data is more reliable following subject replacement.

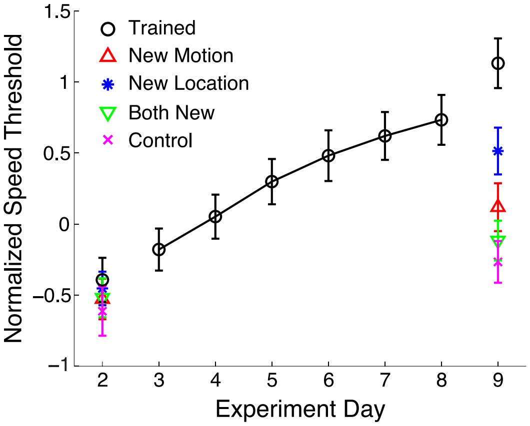

Figure 3. Speed thresholds and learning curve. The trained, new motion, new location, and both new conditions are aggregated across the dot and pinwheel training groups (see Appendix A7 for individual group learning curves). A different staircase procedure was used for the training days (days 3–8) than was used during the two assessment days (days 2 and 9). Error bars represent *SEM*.

Results

For each tracking task, the dependent measure was improvement in tracking speed from baseline, calculated as (posttraining speed − pretraining speed)/ pretraining speed. As shown in Figures 3 and 4, all conditions displayed significant improvement following training (see Appendix A4 for condition means and standard deviations). Considering first only the improvement of participants who completed training sessions (i.e., dot training and pinwheel training subjects), a mixed-factors analysis of variance with within-subjects factor condition (trained, new motion, new location, both new) and between subjects-factor training task (dot training, pinwheel training) revealed a main effect of condition, $F(3, 90) = 21.82$, $p < 0.001$. The main effect of training group, $F(1, 30) = 0.06$, $p = 0.80$, and the interaction of condition x training group, $F(3, 90) = 0.05$, $p = 0.98$, were not significant, indicative of the consistent pattern of improvement between the dot and pinwheel training groups (Figure 4; see Appendix A5 for individual subject distributions).

For both the dot training (dot) and pinwheel training (pw) groups, comparisons between conditions using paired-samples *t* tests revealed that improvement for the trained condition was significantly greater than for the new motion condition, $t_{\text{dot}}(15) = 3.26$, $p_{\text{dot}} = 0.005$, $d_{\text{dot}} = 0.97$; $t_{pw}(15) = 4.11$, $p_{pw} < 0.001$, $d_{pw} = 1.50$); new location condition, $t_{\text{dot}}(15) = 2.51$, $p_{\text{dot}} = 0.02$, $d_{\text{dot}} = 0.87$; $t_{pw}(15) = 4.18$, $p_{pw} < 0.001$, $d_{pw} = 0.88$); and both new condition, $t_{\text{dot}}(15) = 4.20$, $p_{\text{dot}} < 0.001$, $d_{\text{dot}} = 1.28$; $t_{pw}(15) = 5.29$, $p_{pw} < 0.001$, $d_{pw} = 1.78$. Improvement for the new location condition was significantly greater than for the both new condition, $t_{\text{dot}}(15) = 2.26$, $p_{\text{dot}} = 0.04$, $d_{\text{dot}} = 0.59$; $t_{pw}(15) = 2.76$, $p_{pw} = 0.01$, $d_{pw} = 0.84$. Neither the improvement difference between the new location and new motion conditions, $t_{\text{dot}}(15) = 1.18$, $p_{\text{dot}} = 0.26$, $d_{\text{dot}} = 0.27$; $t_{pw}(15) = 1.56$, $p_{pw} = 0.14$, $d_{pw} = 0.49$, nor the new motion and both new conditions, $t_{\text{dot}}(15) = 1.46$, $p_{\text{dot}} = 0.17$, $d_{\text{dot}} = 0.25$; $t_{pw}(15) = 1.86$, $p_{pw} = 0.08$, $d_{pw} = 0.46$, was statistically significant.

To explore whether improvement for the trained, new motion, new location, and both new conditions was greater than would be expected from test–retest related improvement, we compared improvement in each of these conditions to improvement for the no-contact control condition using independent-samples *t* tests. For each control participant, four improvement

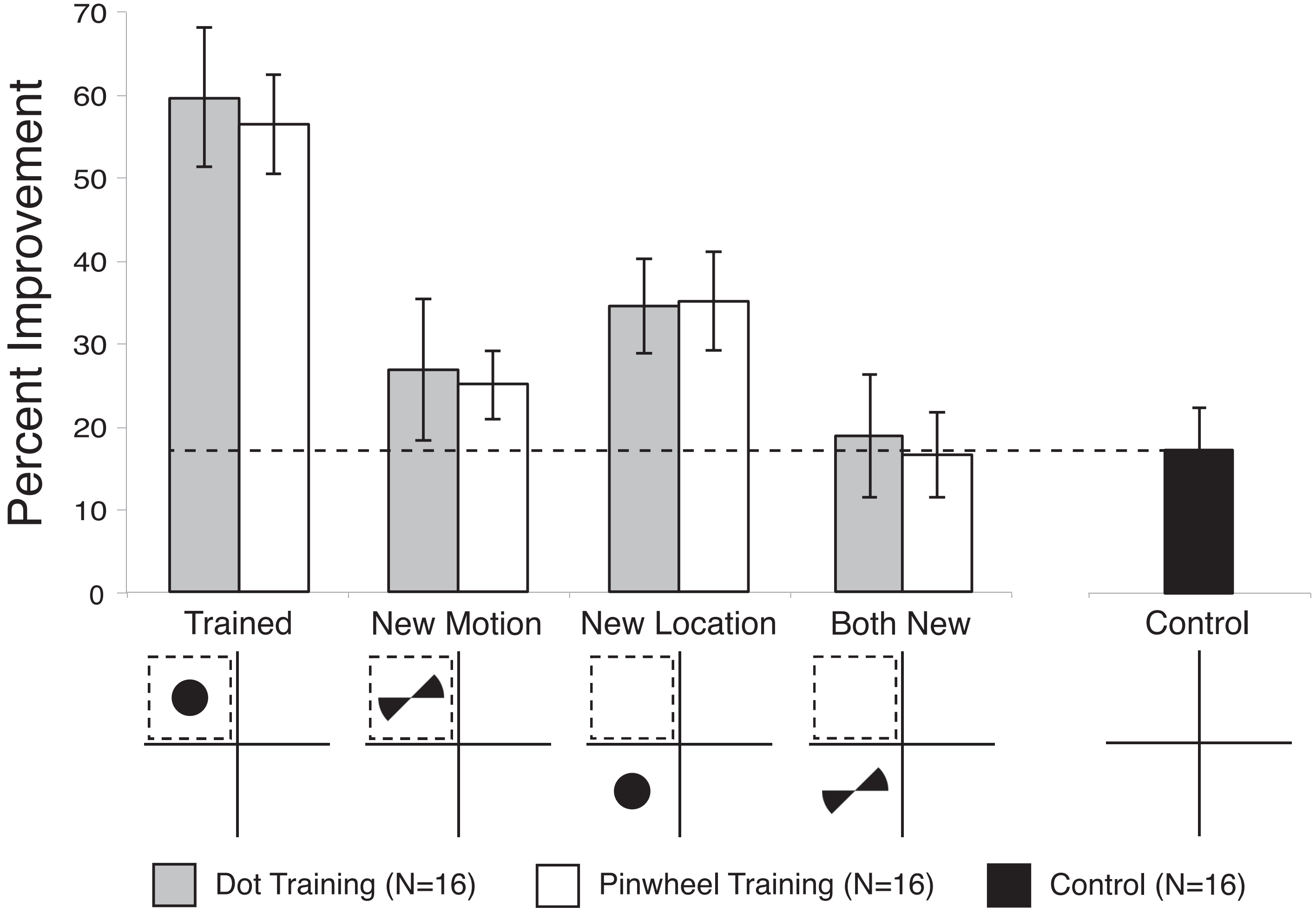

Figure 4. Training results. Dot training and pinwheel training subjects (N = 16 each) trained on their respective training task in one quadrant (shown as dots in the top-left quadrant for illustration here, but counterbalanced between subjects in the actual experiment), whereas control subject (N = 16) completed no training. Speed thresholds were measured before and after training, and the *y* axis shows the percent increase in speed thresholds after training. All tasks were expected to show some improvement due to practice effects. However, the improvement on the trained task in the trained location (Trained) was greater than the improvement when the motion type was different (New Motion), the location was different (New Location), or both were different (Both New) than the training task. Improvement for the trained condition was also greater than for the control condition. The New Location condition showed greater improvement than the Both New condition and the control condition, but no other differences were reliable. Improvement for each condition was significantly greater than 0. Error bars represent *SEM*.

scores were obtained (one for each motion type/location combination). An overall improvement score for each control subject was calculated using the mean improvement in tracking speed across the four motion type/location combinations. Results were again consistent between the dot training and pinwheel training groups. Of primary interest, independent-sample *t* tests revealed that the trained condition showed significantly greater improvement than the control condition, $t_{dot}(30) = 4.28$, $p_{\text{dot}} < 0.001$, $d_{dot} = 1.46$; $t_{pw}(30) = 4.92$, $p_{pw} < 0.001$, $d_{pw} = 1.68$, but that the both new condition did not, $t_{\text{dot}}(30) = 0.15$, $p_{\text{dot}} = 0.88$, $d_{\text{dot}} = 0.05$; $t_{pw}(30) = 0.13$, $p_{pw} = 0.90$, $d_{pw} = -0.04$. Of secondary interest, improvement for the new location condition was greater than that for the control condition, $t_{\text{dot}}(30) = 2.23$, $p_{\text{dot}} = 0.03$, $d_{\text{dot}} = 0.76$; $t_{pw}(30) = 2.26$, $p_{pw} = 0.03$, $d_{pw} = 0.77$, but improvement for the new motion condition was not, $t_{\text{dot}}(30) = 0.95$, $p_{\text{dot}} = 0.35$, $d_{\text{dot}} = 0.32$; $t_{pw}(30) = 1.16$, $p_{pw} = 0.26$, $d_{pw} = 0.40$. See Appendix A6 for model comparisons using mixed-effect analyses, which compare all conditions in a single analysis, and support the same conclusions.

We next tested for baseline differences in standardized threshold speed during the pretraining assessment, in an effort to rule out baseline differences as the cause of the differential training improvement found between conditions. Again considering first only the improvement of participants who completed training sessions (i.e., dot training and pinwheel training subjects), a mixed-factors ANOVA with within-subjects factor condition (trained, new motion, new location, both new) and between subjects-factor training task (dot training, pinwheel training) revealed no main effects of condition, $F(3, 90) = 0.48$, $p = 0.70$, or training task, $F(1, 30) = 0.15$, $p = 0.70$), nor a condition × training task interaction, $F(3, 90) = 1.43$, $p = 0.24$, indicating no reliable differences between training conditions at baseline (Day 2 in Figure 3). We next used independent-samples t tests to test for differences in pretraining speed thresholds between control condition and the other four conditions (aggregated across the dot and pinwheel training groups). These analyses revealed no significant baseline difference between the control condition and the trained ($p = 0.38$), new motion ($p = 0.72$), new location ($p = 0.44$), or both new conditions ($p = 0.68$).

Discussion

The specificity of training benefits found in this study indicates that training did not increase any general attentional capacity, but instead enhanced attention's ability to perform operations over representations specific to the training task, supporting the expertise-building hypothesis of attentional training. Training benefits for a trained motion type in a trained retinotopic location (trained condition) displayed incomplete generalization to tasks differing in seemingly superficial features, such as motion type (new motion), retinotopic location (new location), or both of these characteristics (both new). Furthermore, improvement for the both new condition was no greater than for a control group that performed no training, indicating that no generalization occurred when the properties of motion type and retinotopic location were not shared. Although improvement for the new location condition was significantly less than for the trained condition, improvement for new location was greater than for the both new and control conditions, suggesting that partially shared features between trained and untrained tasks may allow limited generalization. If the large improvement for the trained condition had been the result of enhanced attentional capacity, however, performance gains should have generalized to all the similar untrained tracking conditions.

Since the improvement in the trained condition was not the result of enhanced attentional capacity, at least two alternative possibilities exist. One possibility is that only representations specific to the training task were altered during training (Figure 1), resulting in more precise encoding of target locations. This explanation is compatible with the specificity typically found in perceptual learning studies (Karni & Sagi, 1991; Poggio, Fahle, & Edelman, 1992), where the training task is limited by the quality of representational encoding. If indeed only task-specific representations are enhanced during training, training benefits should generalize to untrained tasks requiring these same representations, even if the untrained tasks require a different attentional capacity. For example, improved tracking of translational motion in the upper-left visual field would be expected to generalize to a search task using the same stimuli, a possibility that could be tested in future studies. However, because the attentional tracking paradigm used here appears to be limited primarily by attentional processing rather than representational quality (Alvarez & Franconeri, 2007; Culham et al., 1998), we favor a second possibility: Training enhanced the coordination between an attentional capacity and representations specific to the training context. That is, rather than improving the quality of perceptual representations specific to the training task, training instead enabled a general attentional capacity to access those representations more effectively. This explanation is compatible with theories positing that enhanced access to representational content is critical for training-induced performance gains (e.g., Ahissar & Hochstein, 2004), and predicts that training benefits will generalize only to tasks requiring attentional expertise that has been enhanced during training (i.e., the connection between the general attentional capacity shared by the tasks and the representations specific to the training task must be enhanced, see red arrow in Figure 1). Therefore, neither a shared attentional capacity nor shared representations between a training and transfer task alone would be sufficient for producing generalization.

Although other studies have reported specificity of training benefits between tasks limited by cognitive control mechanisms, (Gaspar, Neider, Simons, McCarley, & Kramer, 2013; Melby-Lervåg & Hulme, 2013; Thompson et al., 2013), to our knowledge we are the first to demonstrate such minimal generalization of large training benefits between nearly identical attention-limited tasks (all tasks required tracking two moving targets) that were shown to require a common attentional capacity (see Appendix A1). Because of this specificity, we claim that training enhances content-specific attentional coordination rather than general attentional capacities, a conclusion that assumes the attentional tracking tasks in this study use and are limited by the same attentional mechanisms. This assumption can be supported in several ways. First, we ran a preliminary study using the attentional operating characteristic method (see Appendix A1), demonstrating a performance tradeoff between the tasks that is evidence of a common attentional capacity. In previous studies failing to find generalization following training, it is often unclear whether the training and transfer tasks rely on the same capacity [particularly in studies training working memory, which appears to have separate capacities for visual vs. verbal information (Baddeley & Hitch, 1974), visual vs. spatial information (Wood, 2011), and view-dependent vs. view-invariant information (Wood, 2009)]. Second, previous research has suggested that objects tracked within a single visual hemifield (left or right) are maintained by a common attentional resource (Alvarez & Cavanagh, 2005); in the present study, the trained and untrained tasks always occurred within the same visual hemifield. Third, neuroimaging research has implicated the intraparietal sulcus as a critical region across a variety of attentional tracking tasks, including translational (Culham et al., 1998) and rotational motion (Shim, Alvarez, Vickery, & Jiang, 2009); baseline performance for the two types of motion was also highly correlated in our subjects (Appendix A8). Finally, although using target speed as the study's dependent measure (rather than the number of items tracked) may intuitively seem to introduce low-level representational challenges to the tracking task (in addition to attentional limitations), previous research has indicated that observers have sufficient representational precision for tracking items at high speeds, yet are unable to access these representations efficiently when attention is divided (Alvarez & Franconeri, 2007). Furthermore, both the speed at which objects can be tracked and the number of objects that can be tracked appear to be limited by a common capacity (Thompson, Gabrieli, & Alvarez, 2010). Therefore, we are confident that tracking performance was primarily attention-limited in this study despite target speed being the dependent measure, and predict a similar pattern of results would occur if the number of targets were increased during training instead of target speed.

Implications for cognitive training

Inconsistency has plagued training paradigms attempting to enhance the capacity of cognitive control processes critical to everyday functioning, such as attention (Green & Bavelier, 2003) and working memory (Jaeggi et al., 2008). Although we cannot speak directly to the mechanisms underlying generalization in these studies (which often are completed over a longer training period with a more diverse range of training tasks), our results do provide direct evidence against the logic motivating such training paradigms: i.e., that generalization of training benefits is expected between tasks that rely on a common cognitive capacity. Our results indicate that training can produce content-specific attentional expertise, rather than enhancing general attentional capacities; therefore, simply selecting training tasks that share general control mechanisms with outcome measures of interest is insufficient for producing generalization. Until the parameters necessary for producing generalized training benefits are established, consumers interested in cognitive enhancement should be aware that improvement on training tasks is not necessarily evidence of any improvement to general cognitive functioning.

Although the specificity found in the present study challenges the capacity-enhancing hypothesis assumed by many cognitive training studies, we believe that generalization of training benefits is possible under the expertise-building framework we have proposed, particularly if training produces enhanced attentional coordination necessary for the completion of untrained outcome measures. We speculate such generalization may occur when training enhances coordination between distinct control mechanisms (a type of coordination potentially critical to multitasking; Anguera et al., 2013), or when training improves multiple forms of attentional expertise that when combined produce improvement in untrained tasks (e.g., Xiao et al., 2008). Importantly, even training programs that fail to enhance attentional processing may still be beneficial, particularly if learned strategies (e.g., mnemonic training; Verhaeghen, Marcoen, & Goossens, 1992) or placebo effects produce improvement in desirable outcome measures.

Conclusions

In summary, our results suggest that training benefits for attentional tracking are dependent on enhanced coordination of attention with content-specific representations, rather than resulting from an enhanced attentional capacity. We believe the general inconsistency of cognitive training programs is due in part to a misconception that the capacity of cognitive control mechanisms can easily be enhanced in one context for use in another, ignoring the critical importance of content-specific representations in obtaining training benefits. We hope our results will motivate investigation into whether and how enhancing attentional expertise can produce generalized training benefits, potentially leading to more effective and reliable cognitive training paradigms.

Keywords: attention, learning, multiple object tracking, training

Acknowledgments

This study was supported by the National Science Foundation CAREER Award BCS-0953730 to GAA.

RWS was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. The authors thank Elizabeth Tran for assisting with data collection.

Commercial relationships: none.

Corresponding author: Roger Strong.
E-mail: rstrong@fas.harvard.edu.
Address: Department of Psychology, Harvard University, Cambridge, MA, USA.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464.

Allaire, J. C., Bäckman, L., Balota, D. A., Bavelier, D., Bjork, R. A., Bower, G. H., … & Zelinski, E. M. (2014). A Consensus on the Brain Training Industry from the Scientific Community, Max Planck Institute for Human Development and Stanford Center on Longevity. Retrieved from http://longevity3.stanford.edu/blog/2014/10/15/the-consensus-on-the-brain-training-industry-from-the-scientific-community/

Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology*, *18*(3), 337–347.

Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, *16*(8), 637–643.

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13):14, 1–10, doi:10.1167/7.13.14. [PubMed] [Article]

Alvarez, G. A., Horowitz, T. S., Arsenio, H. C., DiMase, J. S., & Wolfe, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(4), 643–667.

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., … & Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature*, *501*(7465), 97–101.

Arend, A. M., & Zimmer, H. D. (2012). Successful training of filtering mechanisms in multiple object tracking does not transfer to filtering mechanisms in a visual working memory task: Behavioral and electrophysiological evidence. *Neuropsychologia*, *50*(10), 2379–2388.

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*(2), 366–377.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89.

Baer, J. (1996). The effects of task-specific divergent-thinking training. *The Journal of Creative Behavior*, *30*(3), 183–187.

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science*, *8*(4), 445–454.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*(7), 349–354.

Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology*, *80*(5), 2657–2670.

Faubert, J. (2013). Professional athletes have extraordinary skills for rapidly learning complex and neutral dynamic visual scenes. *Scientific Reports*, *3*, 1154.

Federal Trade Commission. (2016). Lumosity to pay $2 million to settle FTC deceptive advertising charges for its “brain training” program. *Federal Trade Commission*. Retrieved from https://www.ftc.gov/news-events/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges

Fernandez, A. (2013). Web-based, mobile and biometrics-based technology to assess, monitor and enhance cognition and brain functioning: The digital brain health market, 2012–2020.

Gaspar, J. G., Neider, M. B., Simons, D. J., McCarley, J. S., & Kramer, A. F. (2013). Change detection: Training and transfer. *PloS One*, *8*(6), e67781.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534–537.

Green, C. S., & Bavelier, D. (2006). Enumeration versus multiple object tracking: The case of action video game players. *Cognition*, *101*(1), 217–245.

Harenberg, S., McCaffrey, R., Butz, M., Post, D., Howlett, J., Dorsch, K. D., & Lyster, K. (2016). Can multiple object tracking predict laparoscopic surgical skills? *Journal of Surgical Education*, *73*(3), 386–390.

Holmes, G., & Lister, W. T. (1916). Disturbances of vision from cerebral lesion, with special reference to the cortical representation of the macula. *Brain*, *39*, 34–73.

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.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, *88*(11), 4966–4970.

Legault, I., & Faubert, J. (2012). Perceptual-cognitive training improves biological motion perception: evidence for transferability of training in healthy aging. *Neuroreport*, *23*(8), 469–473.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience*, *3*(12), 1322–1328.

Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis*, *6*(1), 1–20.

Parsons, B., Magill, T., Boucher, A., Zhang, M., Zogbo, K., Bérubé, S., … & Faubert, J. Enhancing Cognitive Function Using Perceptual-Cognitive Training. *Clinical EEG and Neuroscience*, *47*(1), 1–11.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science*, *256*(5059), 1018–1021.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Romeas, T., Guldner, A., & Faubert, J. (2016). 3D-Multiple Object Tracking training task improves passing decision-making accuracy in soccer players. *Psychology of Sport and Exercise*, *22*, 1–9.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., … & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(5212), 889–893.

Shim, W. M., Alvarez, G. A., Vickery, T. J., & Jiang, Y. V. (2009). The number of attentional foci and their precision are dissociated in the posterior parietal cortex. *Cerebral Cortex*, *20*(6), 1341–1349.

Sperling, G., & Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science*, *202*(4365), 315–318.

Thompson, T. W., Gabrieli, J. D., & Alvarez, G. A. (2010). Adaptive training in multiple object tracking expands attentional capacity. *Journal of Vision*, *10*(7): 308, doi:10.1167/10.7.308. [Abstract]

Thompson, T. W., Waskom, M. L., Garel, K. L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., … & Gabrieli, J. D. (2013). Failure of working memory training to enhance cognition or intelligence. *PloS One*, *8*(5), e63614.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging*, *7*(2), 242–251.

Wood, J. N. (2009). Distinct visual working memory systems for view-dependent and view invariant representation. *PLoS One*, *4*(8), e6601.

Wood, J. N. (2011). When do spatial and visual working memory interact? *Attention, Perception, & Psychophysics*, *73*(2), 420–439.

Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, *18*(24), 1922–1926.