Learning to attend: Effects of practice on information selection

Todd A. Kelley
Institute of Cognitive Neuroscience, University College London, London, UK

Steven Yantis
Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA

Though practice can lead to improved performance in many domains, it is currently unknown how practice affects the deployment of selective attention to filter distracting information. We conducted a series of experiments to address this issue by examining how performance on a task changed after repeated exposure to distractors. Distraction initially slowed response time during task performance, an effect that diminished with repeated exposure to the distractors. When the distractors were consistent in appearance, the practice effect developed quickly but was stimulus-specific. When the distractors were more variable in appearance, the practice effect developed slowly but transferred more readily to other conditions. These data indicate that practice with overcoming distraction leads to improvements in information filtering mechanisms that generalize beyond the training regimen when variable distractor stimuli are experienced.

Keywords: attention, vision, practice, distraction

Citation: Kelley, T. A., & Yantis, S. (2009). Learning to attend: Effects of practice on information selection. *Journal of Vision, 9*(7):16, 1–18, http://journalofvision.org/9/7/16/, doi:10.1167/9.7.16.

Introduction

Information selection is a critical component of an organism's successful interaction with the environment. Whether it involves sensory stimuli or internal memories, selective attention to a particular subset of the nearly infinite array of available information is crucial for successfully achieving behavioral goals. This selection process is imperfect and does not always operate under strictly voluntary control. Failures of selective attention give rise to interference with the detection or discrimination of a to-be-attended item, as has been repeatedly demonstrated in the context of visual attention (e.g., Broadbent & Broadbent, 1987; Eriksen & Eriksen, 1974; Folk, Remington & Johnston, 1992; Yantis & Egeth, 1999; Yantis & Jonides, 1984). Although attentional capture by distracting items seems counterproductive, such distraction can be beneficial in alerting the observer to important changes in the environment that might otherwise have been ignored. However, repeated exposure to events that are clearly irrelevant should lead to more efficient filtering of those events. This paper addresses the question of whether and how the deployment of selective attention improves with practice.

Much of the work demonstrating how distracting stimuli interfere with performance comes from studies in which stimulus-driven attentional capture overrides the top-down control signals for directing attention. Several studies by Yantis and colleagues (e.g., Remington, Johnston, & Yantis, 1992; Yantis & Jonides, 1984, 1996) investigated the capture of visuospatial attention by abrupt visual onsets. In these cases, a newly appearing object captures visual attention even when it carries no task-relevant information.

Despite the strength of these attentional capture effects, other studies have shown that attentional capture can be regulated by top-down control. Yantis and Jonides (1990) showed that when subjects were cued to attend to a particular spatial location prior to the onset of a search array, attentional capture by an abrupt onset appearing elsewhere was strongly modulated. Folk et al. (1992) also showed that attentional capture is modulated more generally by the observer's current task set. For example, a to-be-ignored distractor item (e.g., an abrupt onset or color singleton) captures attention only when it matches the visual features of to-be-searched-for target items. This phenomenon of contingent attentional capture (Folk, Leber, & Egeth, 2002; Folk, Remington, & Wright, 1994; Folk & Remington, 1998; Lamy, Leber, & Egeth, 2004; Serences et al., 2005) demonstrates that when subjects direct their attention to specific target features, this task set increases their susceptibility to attentional capture by distractors that contain similar features. Though these studies show that certain cognitive factors modulate distraction, they do not indicate how the susceptibility to distraction might change with practice.

Recently there has been an increase in investigation into how training can improve attention (Tang & Posner, 2009). However, despite the breadth of research addressing questions of attention and distraction, relatively few studies have been concerned explicitly with how learning

doi: 10.1167/9.7.16 Received January 21, 2009; published July 31, 2009 ISSN 1534-7362 © ARVO

changes the effectiveness of visuospatial selective attention. One example, conducted by Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977) showed that when subjects repeatedly searched for letters from a particular target set in rapidly presented arrays, performance gradually reached near-perfect levels. This improvement appeared to be the result of high levels of familiarity leading to an automatic response to the target stimuli, which is supported by the finding that performance decreased to pre-practice levels when the target and distractor letter sets were reversed.

A similar sort of search-related learning has been examined by Chun and colleagues (Chua & Chun, 2003; Chun, 2000; Chun & Jiang, 1998, 1999; Olson & Chun, 2001, 2002; Olson, Chun, & Allsion, 2001), using a paradigm called contextual cueing. In this task, subjects are shown a multiple item search array and are instructed to search for a target item. Trials are divided up into two conditions: one in which the search array is unique to that trial, and one in which the search array is one of a set that has been repeated across multiple blocks of trials. Search performance is measured in response time. In these studies, the difference in performance between old (i.e., repeated) and new search arrays increased with the number of repetitions, with subjects becoming faster at identifying the target in repeated arrays. Furthermore, subjects made use of this learning without being able to explicitly recognize any of the old arrays, indicating that they were implicitly learned. These implicitly learned search arrays apparently came to direct attention to the location of the target, resulting in more rapid target detection than if subjects had searched through the entire array in an unguided fashion. Recent work by Makovski, Vázquez, and Jiang (2008) has expanded on these findings by showing a form of contextual cueing to tracking many moving items simultaneously. In this case, the learned information is restricted to the attended items. In addition, the learning leads to improvements in the dynamic allocation of attention, because the target positions must be tracked across time.

In these examples, the processing of information is guided by a strongly established memory representation, whether it serves to prime a response or to direct attention. This sort of learning can be viewed within the framework of Logan's (1988, 1990) Instance Theory of Automaticity, which holds that task learning occurs through the creation of memory traces that encode the association between a stimulus and its appropriate response. These traces are created through repetition of a particular stimulus on which a particular operation is to be performed (e.g., word/non-word judgment). Initially the operation is performed through the application of relatively slow cognitive operations that may involve visual or memory search, retrieval from long-term memory, and so forth. After many trials, a direct association between the stimulus and the correct response is created in memory. Eventually, recollection of the memory is faster than performance of the cognitive operation(s), and so this comes to drive behavior by bypassing the slower processes.

The studies described thus far principally involve changes in the efficiency with which attention is directed to target items. A related but different question concerns whether the efficiency of selection can itself be changed through learning. Here, the term "efficiency of selection" concerns the degree to which the deployment of selective visual attention to a target item can successfully exclude potentially distracting information from adjacent locations in the scene. Brown and Fera (1994) used a version of the flanker task first discussed by Eriksen and Eriksen (1974) to examine this question. The flanker task provides a direct measure of the efficiency of selective attention: interference by the flanking items indexes a failure to filter them in spite of their task irrelevance. Brown and Fera showed that after a block of practice, the difference in response time to displays containing flankers that were compatible with the appropriate target response and those containing flankers that were incompatible was significantly reduced. This finding does not address, however, whether practice can overcome cases of attentional capture, as opposed to competition at the level of response selection. In addition, this study raises questions about the robustness and generality of attentional practice effects, as Brown and Fera failed to replicate the reduction in flanker interference in a separate experiment with very similar task parameters. Finally, Brown and Fera did not investigate whether practice effects could transfer to previously unseen task or stimulus conditions, leaving open the question of whether improvements in allocation of attention are stimulus-specific or general.

There is evidence in other domains, such as working memory, for practice effects with one task or stimulus set that transfer to other tasks or stimulus sets. For example, Klingberg, Forssberg, and Westerberg (2002) showed that over multiple training sessions, both children with ADHD and healthy adults showed improvements in tasks requiring subjects to remember sequences of items. This training also produced improved performance in untrained working memory tasks as well as general tests of executive control. Olesen, Westerberg, and Klingberg (2004) found similar results with a separate group of subjects. Chen, Eng, and Jiang (2006) reported improvements in a change detection task for novel and familiar abstract shapes after training. More recently, Jaeggi, Buschkuehl, Jonides, and Perrig (2008) showed that practice in demanding working memory tasks produced an improvement in tests of fluid intelligence, which measure the capacity for complex problem solving. These studies show improvements in cognitive operations that generalize well beyond the training context. A major focus of the present studies was to examine this question following improvements in selection efficiency following learning.

The studies we describe here sought to address directly the effects of practice on ignoring distracting stimuli, specifically those that capture visuospatial attention. We asked subjects to engage in a color discrimination task using arrays of red and green dots: on each trial subjects judged whether there were more red dots than green dots, or vice versa. The arrays were generated in such a way that they were unique on every trial, and so prevented learning of stimulus-response associations. Highly salient abrupt onset distractors consisting of stimuli from many different categories were used, and their effects on response time (RT) were measured. Improvements in the efficiency of selective attention were quantified by the change in the effect of distractors over time. We investigated the generality of these improvements by examining transfer of practice across previously unobserved distractor conditions. The results reveal that practice leads to improvements in ignoring distracting items and that these improvements are more general when the distractors have greater variability in their appearance.

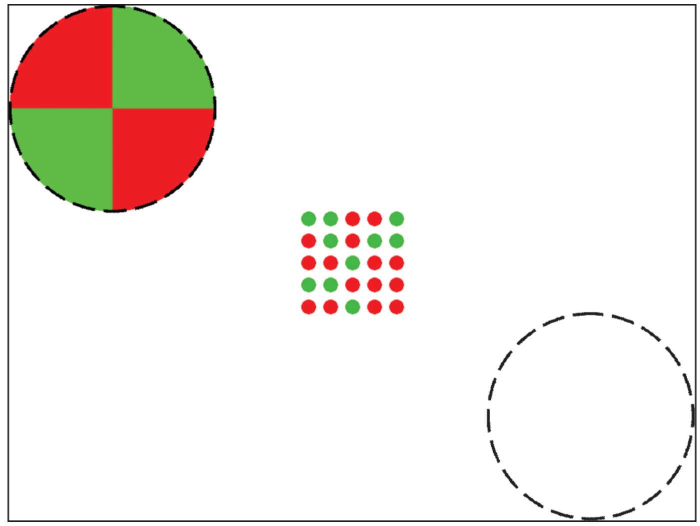

Figure 1. Stimuli used in Experiment 1. Subjects were asked to determine whether more red or green dots were present in the 5 × 5 array at the center of the screen. The array was presented for 100 ms. On half of all trials, a large red and green distractor (shown in upper left corner) was presented in one of the two locations indicated by dashed circles (not actually present in the experiment). The distractor appeared 100 ms before the array onset and stayed on the screen until the task array disappeared.

Experiment 1

Methods

Participants

Twenty Johns Hopkins University undergraduates (10 males, aged 18–21) enrolled in Psychology courses participated in this experiment for partial course credit. Subjects had normal or corrected-to-normal visual acuity and normal color vision. All subjects provided informed consent approved by the Johns Hopkins University Institutional Review Board.

Stimuli

Stimuli were generated using custom MATLAB software (MathWorks, Natick, MA) running the Psychophysics Toolbox (v 2.54; Brainard, 1997; Pelli, 1997). Subjects viewed an LCD display screen from a distance of approximately 40 cm.

Stimuli consisted of a 5 × 5 array of dots, centered in the display (Figure 1). Each dot was 0.6 degrees of visual angle in diameter, and the edge-to-edge distance between dots was 0.375 degrees; the entire array comprised a square subtending 4.5 by 4.5 degrees. Each dot was either red (RGB value: 255 0 0) or green (RGB value: 0 225 0); there were always 15 dots of one color and 10 of the other. The dots were positioned randomly within the array, with the following constraints: at least two dots of each color had to be present across both the leftmost and rightmost two columns, as well as in the topmost and bottommost two rows (i.e., of the ten dots in the two leftmost columns, no more than eight could be the same color; this was also true of the two rightmost columns and the top and bottom two rows). These constraints served to prevent the dots from being grouped in such a way that determining which color was in the majority became trivial.

Distractor disks were also presented on some trials, as shown in Figure 1. These disks had a diameter of 9 degrees and were made up of alternating red and green quadrants. The distractors were placed in either the upper left or lower right quadrants of the display and positioned such that the center of the disk was 16.4 degrees from the center of the dot array. While a distractor was onscreen, the red and green quadrants reversed color at a rate of 10 Hz, so that the disk appeared to be flickering or rotating. All stimuli were presented on a white background.

Procedure

Subjects were given written instructions at the beginning of the session; an experimenter was also present at this time to answer any questions. These instructions explained that the subject would see a briefly presented array of dots on each trial and that they were to determine whether there were more red or green dots in the array. They responded by pressing the J key for red and the N key for green. Responses were made with the subject's dominant hand. Subjects were also told that if they responded incorrectly, or took too long to respond, they would hear a tone.

Following the instructions, subjects completed a practice block of 50 trials. Each trial consisted of the array being presented for 100 ms, followed by 2.4-s inter-trial interval. If the subject did not respond within the allotted time, the response to that trial was recorded as incorrect. No distractors were presented during the practice block. The experimenter observed the subject during this practice period to ensure that they understood the task.

The main task consisted of 10 blocks; each block contained 100 trials, a random half of which contained distractors. The distractor appeared equally often in the upper left or lower right locations in a random order. These two factors were counterbalanced with majority array color (equally often red and green). On trials that contained a distractor, the distractor appeared 100 ms before the array and stayed on the screen for the duration of the array presentation, for a total duration of 200 ms. Subjects were explicitly informed that there would be a distractor on some trials and that it should be ignored.

To encourage both speed and accuracy, subjects' performance was scored throughout the main task. Each trial was worth a base value of 100 points; this value decayed by 1 point every 10 ms that elapsed until the response, starting at the offset of the array. If a correct response was made, the subject received the remaining point value for that trial (this value could never be less than zero for a correct response). If an incorrect response was made, a 200-point penalty was assessed and subtracted from the subject's cumulative point total. The scoring system was explained to the subjects during the instruction period. They were also told that, even though speed and accuracy were both important, they would lose more points from errors than they could gain by making rapid but haphazard responses. At the end of each block, subjects were shown a breakdown of their performance for that block, including their average response time (RT, measured from the array onset) and the points subsequently awarded for those correct trials, incorrect responses and the total penalties assessed as a result, and their total score for the block, along with their score so far for the experiment. Subjects were encouraged to earn as high a score as possible, and if they scored fewer than 1000 points for the block (corresponding roughly to 90% accuracy and a mean RT of 770 ms), they received a message that said "Please try to do better." A single session lasted approximately 50 min.

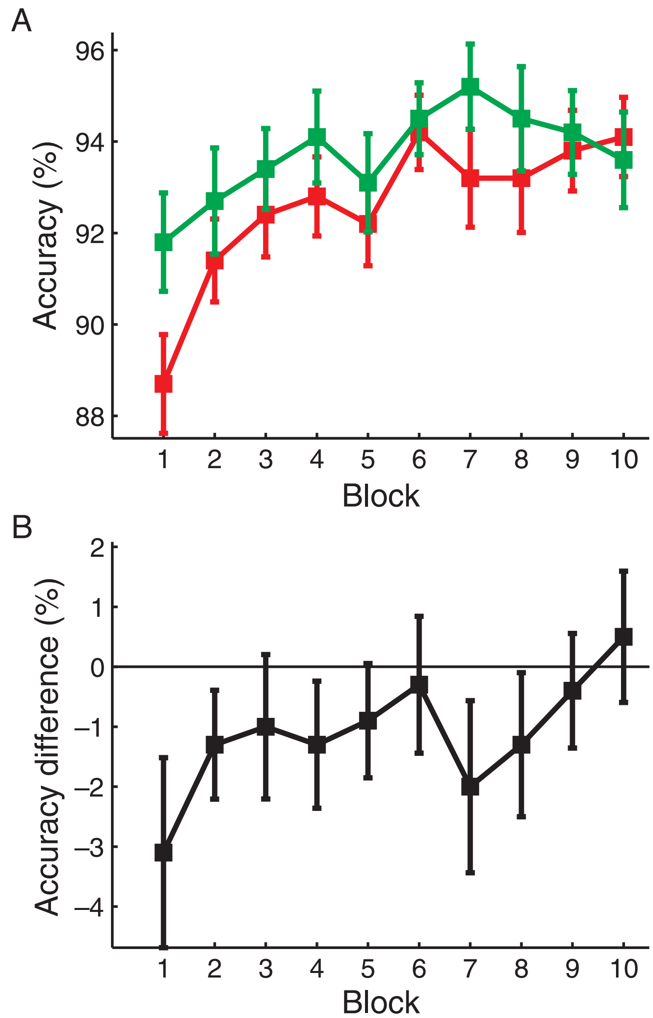

Figure 2. (Top) Mean % Correct for Distractor Trials (red line) and No Distractor Trials (green line) as a function of block. (Bottom) Difference in task accuracy (Mean % Correct for Distractor Trials − Mean % Correct for No Distractor Trials) as a function of block, showing that the presence of distractors had an initial impact on task performance, but that this effect declined with practice. Error bars on all plots represent ±1 *SEM*.

Results

Mean accuracy for the task was 93.3% (No Distractor trials: 93.7%; Distractor Trials: 92.6%). Mean RT was 733 ms (No Distractor: 730 ms; Distractor: 735 ms). Subjects were both slower and less accurate when a distractor was present than when one was not present, indicating the presence of distractors had an overall negative impact on performance.

Figure 2 shows the effects of distraction on accuracy as a function of practice. These two plots show the difference in accuracy between Distractor and No Distractor trials, averaged across all subjects for each experimental block (negative values indicate that performance was worse when a distractor was present). Paired *t*-tests revealed a trend for accuracy on block 1 to be worse for Distractor than No Distractor trials ($t(19) = 1.91$, $p = 0.072$); no such trends were present for any other block. Though not statistically reliable, the accuracy data follow the same qualitative pattern as the RT data, discussed below. This was the case for all the experiments to be discussed here; therefore, further analysis will focus on the effects of distraction on response time.

The effect of distraction on RT was quantified as follows. For each subject, Distractor and No Distractor trials were paired based on order of presentation (i.e., Distractor Trial 1 was paired with No Distractor Trial 1, Trial 2 with Trial 2, etc.). Because there were an equal number of distractor and no distractor trials in each block, trials were never paired across different blocks. The difference in RT was calculated for each trial pair.

The bottom plot of Figure 3 shows the RT difference as a function of experimental block: each point represents the mean difference for all trial pairs within a block, averaged across all subjects (positive values indicate that responses were slower when a distractor was present). This plot shows that RT was slower when distractors were present than when they were absent, but only for the first block ($t(19) = 3.53$, $p < 0.002$; for all other blocks, $p > 0.05$); the ANOVA for the effect of block on RT difference was not significant ($F(9,171) = 1.66$, $p = 0.1$).

Discussion

Subjects were slower to make color discrimination judgments when an abrupt-onset distractor appeared just before the task array than when it was absent. After approximately 100 trials, subjects were able to successfully filter the distracting information, eliminating the RT cost. Thus, practice led to an improvement in the efficiency of selective attention.

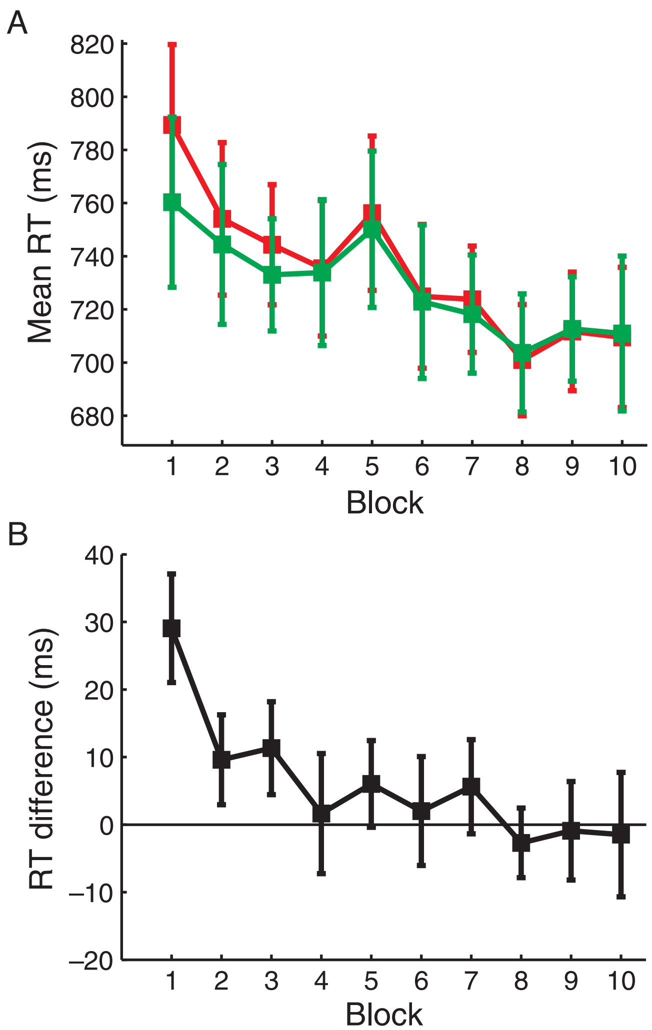

Figure 3. (Top) Mean RT for Distractor Trials (red line) and No Distractor Trials (green line) as a function of block. (Bottom) Difference in response time (Distractor Trial RT − No Distractor Trial RT) as a function of block, showing that distractors initially produced slower RTs, but that this slowing declined with practice.

These results raise the question of whether the learning was highly specific to the distractors to which subjects were exposed in this task, or whether it was a general improvement in selective efficiency that would apply to any distracting stimulus. If, after learning, the distractors changed location, shape, or manner of appearance, would their capacity to distract attention and hence slow responses return? The extent to which practice carries over to new task conditions is dependent on the means by which distractor suppression is taking place. One possibility is that subjects were able to focus their attention on the locus of the array at the start of each trial. Yantis and Jonides (1990) showed that when attention is explicitly cued to a specific region of space, abrupt onset items that appeared elsewhere failed to capture attention. In the present experiment, subjects were not specifically cued to the location of the array prior to its onset. However, the location was stable throughout the experiment, always appearing at the center of the display. Subjects might have used this knowledge to improve their focus of attention to the center of the display, and so prevent attentional capture by the distractors.

Another possibility is that the reduced distractor effect is the result of visual adaptation. Studies have shown that repeated exposure to a stimulus produces a reduction in the neural response to that stimulus, both at the level of single cells and populations of neurons (Grill-Spector, Henson, & Martin, 2006; Krekelberg, Boynton, & van Wezel, 2006). The repeated presentation of the distracting elements could have reduced the response to the distractors, causing them to be less effective qua distractors. This possibility will be addressed in later experiments.

Spatially specific inhibition of any sensory information appearing in the distractor locations due to top-down modulatory signals could also reduce distraction (Awh, Matsukura, & Serences, 2003; Serences, Yantis, Culberson, & Awh, 2004). Such inhibition would likely take time to be established, as there would be initial uncertainty as to the nature, location, and task relevance of the distractors. Once established, such suppression could be an effective means of preventing distraction.

The possibilities described above lead to different predictions regarding transfer of practice effects. If practice effects are the result of improvements in directing attention to the array location on every trial, then one would expect transfer of these practice effects to new distractor conditions, such as a shift to new locations or different distracting items. However, if sensory adaptation produced the reduction in distraction, then there should be no transfer to new distractor conditions, as adaptation is generally tied to the precise sensory characteristics of the adapting stimulus (i.e., location and features). Similarly, if subjects are exercising improved suppression at the distractor locations, then practice effects should only transfer to new distractor items and not to new locations. To examine these possibilities, the next two experiments investigate transfer of practice effects to new distractor locations and to new distractor items.

Experiment 2

Methods

Methods were the same as those described in Experiment 1, with the following exceptions.

Participants

Thirty-one Johns Hopkins undergraduates (12 males, aged 17–24) participated in this experiment for extra course credit. Data from one subject were excluded from the final analysis because of irregularities encountered during the session.

Stimuli and procedure

For all participants, the distractor items appeared in one of two quadrants for the first five blocks of the task (for one group, the initial positions were the upper left and lower right quadrants, and for the other group, the other two quadrants). After five blocks of the task, the position of the distractors switched to the lower left and upper right for the first group, upper left and lower right for the second group. Subjects were not informed in advance that this change would occur.

Results

The mean accuracy across all subjects for the main task was 91.6% (No Distractor: 91.9%; Distractor: 91.2%). The mean RT for correct trials was 756 ms (No Distractor: 752 ms; Distractor: 760 ms). The RT difference for Distractor and No Distractor trials was calculated as in Experiment 1. A mixed factorial ANOVA examining RT difference as a function of the within-subjects factor Block and the between-subjects factor Initial Distractor Location revealed no main effect of Initial Location ($F(1,28) = 1.16$, $p > 0.2$), nor a significant interaction ($F < 1$); therefore all further results will collapse across the two groups of subjects.

Figure 4 shows the RT difference as a function of block. As in Experiment 1, there was an initial slowing of RT due to the presence of distractors; this effect declined quickly with practice. The plot also shows that, following the change of the distractor locations between blocks 5 and 6, the distractor effect returned, indicating that the appearance of distractors in new locations reinstated the distraction effect that was present before any practice with the task. Following the pattern in the first half of the experiment, the renewed distractor effect was quickly reduced with further practice. To examine how the change in distractor location affected practice-related changes to the RT difference, a within-subjects ANOVA was carried out on the factors Session Half (first half (old locations) and second half (new locations)) and Block (Blocks 1–5 within each session half). There was a trend toward a significant effect of Block ($F(4,116) = 2.36$, $p = 0.058$), but no significant effect of Session Half, and no significant interaction (Fs < 1). T-tests showed that both Block 1 ($t(29) = 2.02$, $p = 0.05$) and Block 6 ($t(29) = 2.28$, $p < 0.03$) yielded RT differences significantly greater than zero. Furthermore, the RT differences for Blocks 1 and 6 were not significantly different from each other ($t < 1$). This suggests that there was no transfer of the practice effects when the new distractor locations were introduced.

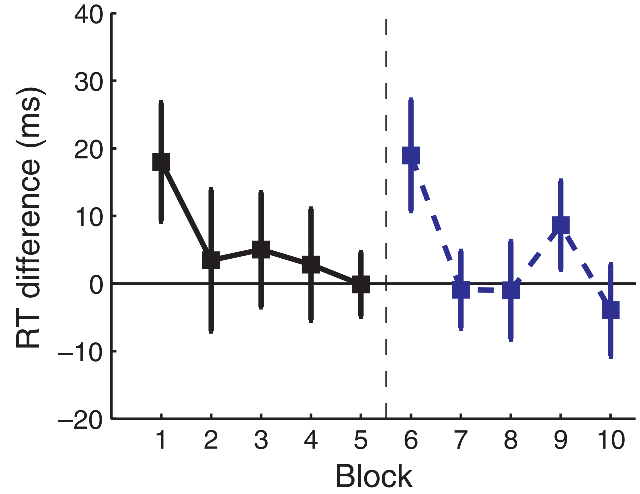

Figure 4. RT Difference between Distractor and No Distractor conditions as a function of Block. Vertical dashed line indicates the point at which distractor locations were switched (between Blocks 5 and 6). Solid black line: data for old distractor locations; dashed blue line: data for new distractor locations.

The trial pairing procedure described in Experiment 1 allowed us to examine the RT difference at a more fine-grained level, allowing for the observation of practice effects within blocks. Figure 5 shows the effect of practice within Blocks 1 and 6, respectively. Each point depicts the RT difference for one bin of 5 trial pairs in Block 1 (solid line) and Block 6 (dashed line). This plot shows that, on a trial-by-trial level, there appear to be some differences in the magnitude of the distractor effect between the two blocks. A Block (1 vs. 6) by Bin ANOVA showed that there was no significant effect of Block ($F < 1$), but there was a significant effect of Bin ($F(9,243) = 2.23$, $p < 0.02$) and a significant interaction between the two conditions ($F(9,243) = 2.16$, $p < 0.03$). A post-hoc contrast comparing

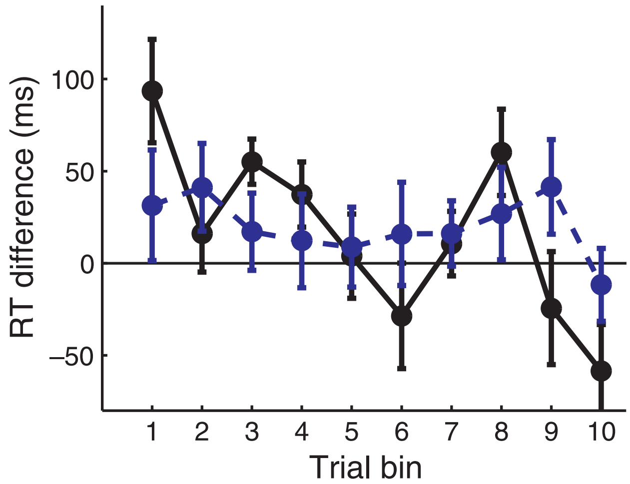

Figure 5. RT difference as a function of trial bin within Blocks 1 and 6, respectively. Solid black line: Block 1 (old distractor locations); dashed blue line: Block 6 (new distractor locations). Each bin contains five pairs of Distractor and No Distractor trials; the plot shows the mean RT difference of all trial pairs in a bin, averaged across all subjects. Note the increase in scale used here to accommodate the larger range of RT differences, compared to previous figures.

the difference in the linear trend between Block 1 and Block 6 was also significant ($F(1,243) = 5.68$, $p < 0.03$), indicating that the drop off in RT difference in Block 6 had a shallower slope than in Block 1. The relatively flat function for RT difference in Block 6 suggests that distractors did not have the same impact in Block 6 as they did in Block 1.

Discussion

The results of Experiment 2 replicate those of Experiment 1, showing that with repeated exposure to a distracting stimulus, subjects were able to improve their capacity to ignore this distracting information, leading to a reduced RT cost. This experiment also revealed that when the distractors appeared in new locations, there was little carryover of the practice effects. The mean RT cost was the same at the start of the experiment (Block 1) as it was when the new locations were introduced (Block 6). However, the temporal pattern of this RT difference was itself different within those two blocks.

Though it is tempting to suggest that the shallower slope of RT difference as a function of time in Block 6 is indicative of training transfer, such an interpretation is problematic, especially since the mean values across the two blocks were not different. The most clearly recognizable case of transfer would have been for the data in Block 6 to have a shallower slope in conjunction with a mean value that is consistently lower than that of Block 1. Though this did not occur, it is possible that the observed data could be the result of some carryover of information, depending on the origins of the practice effect.

The lack of clear transfer effects rule out an explanation of the initial results based on improvements only in the efficiency with which the target location is selected. Throughout the experiment, the location of the main task array was consistent. If subjects simply improved the efficiency of spatial attention to the target location, thus avoiding capture by abrupt onsets, this should have led to continued efficient selection following the location switch. Thus, subjects' attention would continue to be cued to the location of the array, and the new spatial locations of the distractors should have had no impact on performance. The data rule out this possibility.

The results are consistent, however, with an account in which spatially specific suppression of the distractor locations is learned through practice. In this condition, the change in location would have disrupted such a location-specific improvement, leading to a period of relearning to suppress distractors in the new locations. If the observed practice effects were the result of location suppression alone, then a change in the distractor identity while maintaining the same location should yield strong transfer of practice effects.

The next experiment will examine transfer of practice effects when the distractor identity changes (i.e., from the red and green disks to pictures of faces), but the location remains the same. If transfer is observed in this condition, it would offer further evidence for location-specific distractor mechanism. If no transfer is observed, it would suggest a much more stimulus-specific means of distractor filtering.

Experiment 3

Methods

Methods were the same as those described in Experiments 1 and 2, with the following exceptions.

Participants

Thirty-four Johns Hopkins undergraduates (12 males, aged 18–24) participated in this study for extra course credit. Data from two subjects were excluded from the final analysis because of poor behavioral performance (mean accuracy less than 80% for the session).

Stimuli and procedure

In this experiment, grayscale pictures of faces (taken from Hemera Photo-Objects 50,000, vol. II by Dundas Software, 2000; see Experiment 4 and Figure 8, for an example) were used as the distractors for the first half of the experiment for half of the subjects. These pictures were scaled to fit roughly within the same space in the display subtended by the red and green distractor disks. There were forty face pictures used in total, presented in such a way that no picture was seen more than twice per block (40 faces vs. 50 distractor trials per block). After five blocks of the main task, the distractor type switched (from disks to faces for the first group, and from faces to disks for the second group) and stayed this way for the remaining five blocks of the task. The instructions regarding the presence of distractors was the same as in the previous experiments: subjects were told that throughout the session, objects other than the dot array may appear on the screen, but that they were to ignore these objects and perform the task as well as possible. They were never told of the potential identity of these distractor objects, nor were they told that the identity might switch.

Results

The mean accuracy across all subjects was 91.9% (No Distractor: 92%; Distractor: 91.8%). The mean RT was 705 ms (No Distractor: 704 ms; Distractor: 705 ms). RT difference was calculated as in the previous experiments. A mixed factorial ANOVA examining the within-subjects factor of Block and the between-subjects factor of Initial Distractor Item revealed no significant effect of Initial Item ($F < 1$), nor a significant interaction ($F(9,270) = 1.35$, $p > 0.2$); therefore, all further results will collapse across the two groups of subjects.

Figure 6 shows RT difference as a function of Block, averaged across all subjects. This plot shows roughly the same pattern as in Experiment 2, specifically an initial (though somewhat smaller) effect of distractor that quickly dropped off, followed by a return of the distractor effect when the new items were introduced. A Session Half by Block ANOVA showed a significant effect of Block ($F(4,124) = 5.79$, $p < 0.001$), but no significant effect of Session Half, and no significant interaction (Fs < 1). *T*-tests showed that only the RT difference for Block 6 (introduction of new distractor item) was significantly different from zero ($t(31) = 3.27$, $p < 0.003$). A paired *t*-test showed that the RT difference for Blocks 1 and 6 were not significantly different from each other ($t(31) = 1.63$, $p > 0.1$).

Figure 7 shows the RT difference for Blocks 1 (solid line) and 6 (dashed line), plotted against bins of 5 trial pairs. This plot shows that the initial effect of distractor was not substantially lower than in the previous experiments but rather declined to zero within the first few trials, explaining why the RT difference averaged across Block 1 was not significantly different from zero. Unlike Experiment 2, the results for Blocks 1 and 6 were similar, with Block 1 having a somewhat lower overall RT difference than Block 6. A factorial ANOVA examining the effects of block and trial bin showed a significant effect of Block ($F(1,31) = 5.24$, $p < 0.03$) and Bin ($F(9,279) = 3.77$, $p < 0.001$) but no significant interaction ($F < 1$). The pattern of results observed here indicates that practice effects did not survive the change in distractor item.

Discussion

Experiment 3 examined the capacity for transfer of practice effects to new distractor items. In contrast to the weak evidence for some degree of transfer seen in Experiment 2, here there was no evidence of any transfer

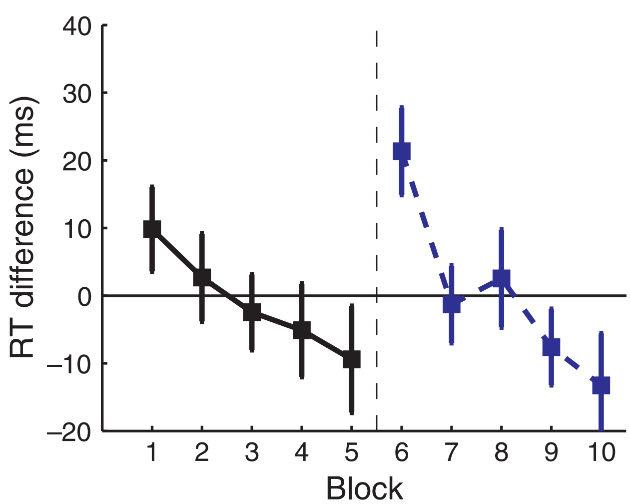

Figure 6. RT difference between Distractor and No Distractor trials as a function of Block. The vertical dashed line indicates the point at which distractor items were switched (between Blocks 5 and 6). Solid black line: old distractor items; dashed blue line: new distractor items.

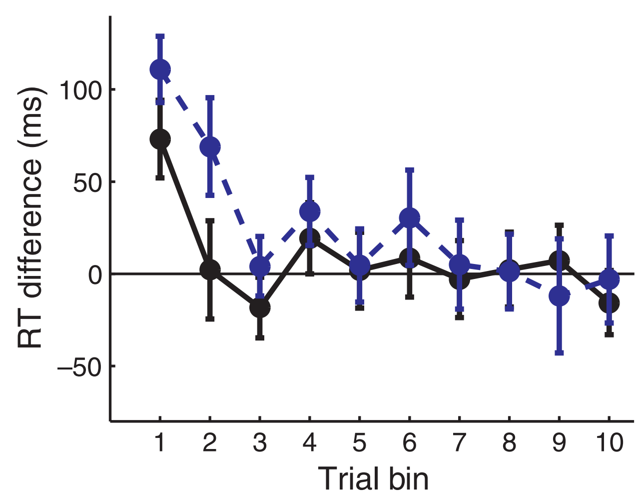

Figure 7. RT difference as a function of trial bin. Solid black line: Block 1 (old distractor items); dashed blue line: Block 6 (new distractor items). Note the difference in scale compared to Figure 6.

following a change in distractor. If anything, subjects appeared to take longer to suppress the new distractor items than the ones presented at the outset of the session. This effect seems to be the result of subjects who were highly efficient at learning to suppress distractors (Figure 7, Block 1 data), rather than any experimental differences between this study and previous studies.

These data further constrain the possible explanations for the observed results thus far. The lack of any transfer effects strongly argues against a location-specific suppression. If subjects were effective at filtering sensory information based on location, then the change in identity of those distractors should have had no impact on performance. Even a conjunction between item-based and location-based filtering would have produced some transfer to the new distractor items, in contrast to the observed data.

These data also cast doubt on a sensory adaptation explanation. As stated above, there was no difference in the performance of subjects who saw faces first and that of subjects who saw disks first. This indicates that the improvement in selective efficiency was the same for both types of distractors. However, as indicated in the Methods section, there was much less repetition for the individual face stimuli (no more than two presentations per block, and only for 25% of the faces). This excludes the possibility that participants showed adaptation to individual face images, as practice effects were observed within a single block. Although some fMRI studies have shown adaptation to face stimuli in general (as opposed to individual faces; e.g., Furl, van Rijsbergen, Teves, & Dolan, 2007; Webster, Kaping, Mizokami, & Duhamel, 2004), those adaptation effects were for features such as gender, race, and expression; the faces used here were not uniform in any of those categories, preventing adaptation to any of those features. In addition, such face adaptation paradigms tend to use much longer presentation times than those used in these experiments, making the current distractor presentation methods unsuitable for sensory adaptation.

Gati and Ben-Shakhar (1990) also showed habituation of the orienting response (OR) for individual features of faces, using composite illustrations that could be formed by combining various facial features onto a single template. This raises the possibility that subjects habituated to individual features presented across multiple faces. Though some faces within the group may have shared certain features (hair style, nose size, etc.), this would only have applied to certain subsets of the face images, rather than the entire set. Furthermore, there would have greater variance in appearance than in the images used by Gati and Ben-Shakhar, because the images were of real people. As such, face adaptation is unlikely to account for the full reduction in the distractor effect observed here (see General discussion section for more on habituation of the OR).

Of the previously discussed explanations, a filtering mechanism that is tied to inhibition of the specific distractor stimuli at the specific locations is most consistent with the results of Experiments 1–3. Given the highly consistent nature of the distractors (flashing red and green disks or similarly shaped grayscale faces, all appearing in one of two possible locations), very specific stimulus inhibition mechanisms for filtering distracting information (e.g., inhibition of red and green circular shapes at specific locations in the visual field) would have been effective. Furthermore, the consistency of the distractors would enable such a mechanism to develop quickly. If the distractors were more heterogeneous in appearance, then a more general-purpose selection mechanism might be required. Such a strategy could produce different results, especially with regard to transfer. Experiment 4 will address this possibility by substantially increasing the heterogeneity of distractor stimuli used.

Experiment 4

To examine whether the heterogeneity of distractor items influences the nature of improved selection efficiency, we used distractors from many different object categories in Experiment 4. Along with the new set of stimuli, two more changes were introduced in this experiment. First, all subjects participated in two sessions of the experiment. In the first session, only the main task was presented, without distracting items. This allowed subjects to reach asymptotic performance on the primary task and permitted investigation of whether or not high levels of task familiarity would speed the learning process. Second, both new and old distractor locations were used in post-transfer blocks. This allowed for a direct comparison between the old and new distractor conditions.

Methods

Methods were the same as those described previously, with the following exceptions.

Participants

Thirty-two people (12 males, aged 18–54) recruited from the Johns Hopkins University community participated in two sessions of this experiment for monetary compensation. Subjects were paid a base rate of $8/session, plus a bonus based on their score. Data from two subjects were excluded from the final analysis: one because of poor performance (mean accuracy less than 80% for one of the sessions) and one because data from one of the cells in the RT difference analysis was more than three standard deviations away from the group mean; this did not change the overall pattern of the results.

Stimuli and procedure

Subjects were recruited to participate in two sessions. During the first session, they completed one practice block and ten blocks of the main task; no distractors were presented during this session. During the second session, subjects completed the same number of trials, and distractors were presented on a randomly selected half of the trials (excluding practice trials). Distractors consisted of a set of 520 distinct images belonging to various classes of objects. These classes included color and grayscale picture of faces, grayscale pictures of hand tools, furniture and animals, color illustrations of familiar objects, and computer generated color fractal patterns (see Acknowledgements for image sources). Each picture was scaled to fit within a square measuring 9 degrees of visual angle on a side. The distractors were positioned closer to the dot array than in previous studies, with a center-to-center distance of 10.1 degrees between the distractor and the array. One distractor appeared on each distractor-present trial (see Figure 8).

During the first five blocks of the second session, the distractor appeared in either the upper left or lower right positions. For the last five blocks, the distractor appeared equally often in any of the four positions around the dot array. Each of these five blocks contained 104 trials (52 distractor-present and 52 distractor-absent). Each of the 520 distractor images was seen only once by each subject. Subjects' performance was scored using the same procedure as in the previous experiments. Subjects were paid a bonus based on their score, at a rate of \$1 per 10,000 points in each session. Subjects were informed of the bonus system at the start of the first session.

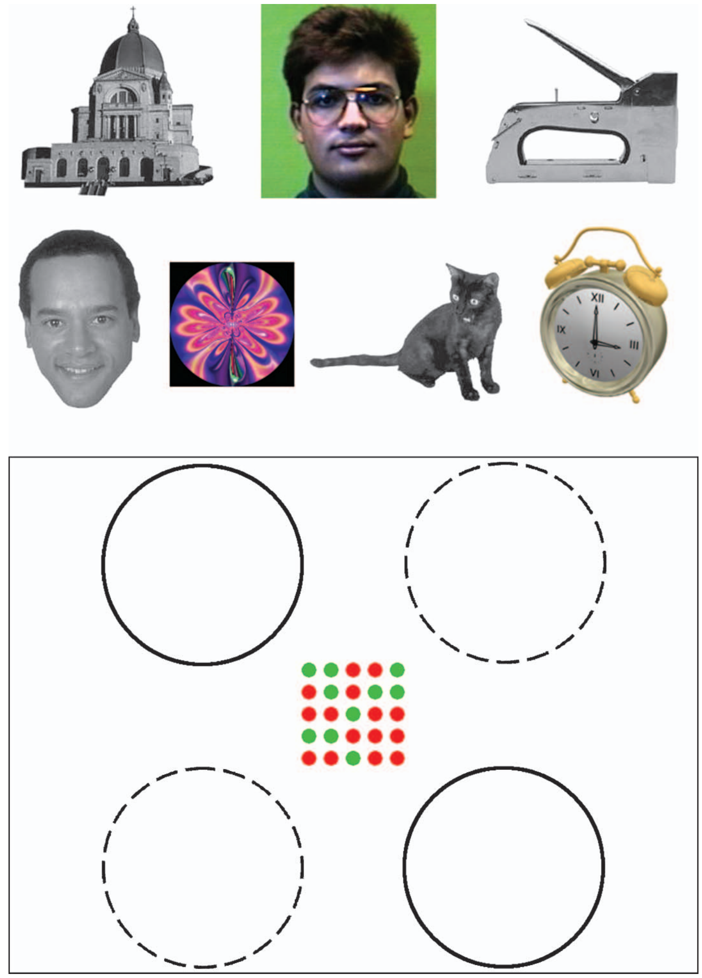

Figure 8. Distractor items and locations for Experiment 4. (Top) Examples from each class of image used as distractors: grayscale buildings, color faces, grayscale objects, grayscale faces, color fractals, grayscale animals, and color objects. (Bottom) Distractor locations (lines not visible during experiment). Solid lines indicate initial locations used throughout the experiment. Dashed lines indicate new locations introduced at the beginning of Block 6.

Results

Session 1—no distractors

The mean accuracy across all subjects for the first session was 94%; the mean RT was 735 ms. Subjects become faster and more accurate on the main task with practice. Practice (as measured by Block) had significant effects on RT ($F(9,261) = 5.39$, $p < 0.001$) and accuracy ($F(9,261) = 2.17$, $p < 0.03$).

Session 2

The mean accuracy across all subjects was 93.9% (No Distractor: 94%; Distractor: 93.8%). The mean RT was 719 ms (No Distractor: 706 ms; Distractor: 733 ms). A paired *t*-test focusing on non-distractor trials showed an overall improvement in task RT for Session 2 compared to Session 1 ($t(29) = 3.03$, $p < 0.005$). However, this improvement did not make the subjects in Experiment 4 any faster on non-distractor trials than subjects from the previous experiments ($F(3,108) = 1.33$, $p > 0.27$; $p > 0.29$ for Tukey's HSD comparing Experiment 4 RTs individually to Experiments 1–3).

Figure 9 shows RT difference as a function of block for both old distractor locations (solid line) and new distractor locations (dashed line, Blocks 6–10). As can be seen by this plot, subjects took longer to overcome the distractor effect than in previous experiments. In addition, comparing the solid and the dashed lines for Blocks 6 through 10, it is clear that there was no difference in performance when distractors appeared in the new locations, compared to when they appeared in the old locations. Examining old and new locations separately for Blocks 6 through 10 showed that effects of Location, Block, and their interaction were not significant (all *F*s < 1). Collapsing across old and new locations, a Session Half by Block ANOVA showed a significant effect of Session Half ($F(1,29) = 16.13$, $p < 0.001$) but no significant effect of Block ($F(4,116) = 1.97$, $p < 0.11$). There was a significant

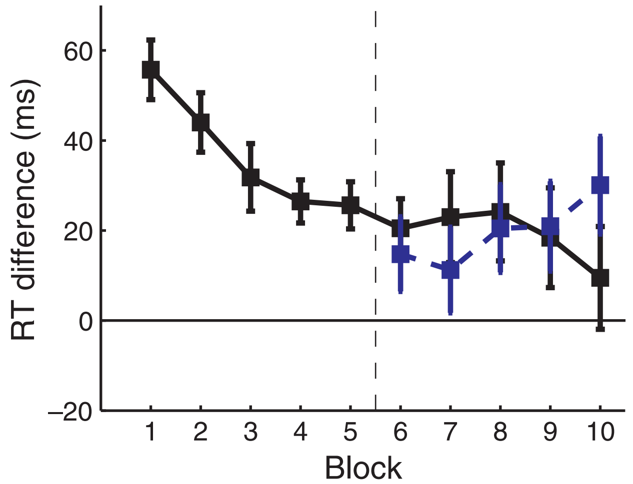

Figure 9. RT difference between Distractor and No Distractor trials as a function of Block. Vertical dashed line indicates point at which new locations were introduced (between Blocks 5 and 6). Solid black line: data for old distractor locations; dashed blue line: data for new distractor locations. NB: the scale for the RT difference is expanded compared to previous plots.

interaction between these two factors ($F(4,116) = 2.94$, $p < 0.025$). This interaction was well explained by the change in linear trend across Blocks from the first half to the second half of the session ($F(1,29) = 9.95$, $p < 0.005$), which accounted for 87% of the variance due to the interaction. *T*-tests indicated that the RT difference was greater than zero for all blocks ($t_{\text{Block } 7}(29) = 2.31$, $p < 0.03$; $t_{\text{Block } 9}(29) = 2.60$, $p < 0.02$; $p < 0.01$ for all other blocks).

Figure 10 shows the RT difference for Block 1 and the old and new locations for Block 6 as a function of trial bin. Given that there were 26 trials for each distractor location in Block 6, the trials were binned differently than previously described: Bins 1 and 3 had seven trial pairs and Bins 2 and 4 had six trial pairs for both old and new locations in Block 6 (the bins for Block 1 had twice the number of trial pairs per bin). This plot further shows the transfer of practice effects to the new distractor locations. Block 1 (solid black line) showed a substantial distractor effect that took longer to approach zero; both old (dashed black line) and new (blue line) locations in Block 6 showed a lower distractor effect and flatter slopes. Additionally, there was extensive overlap between the old and new locations in Block 6. The effect of Block (Block 1 vs. Block 6 old vs. Block 6 new) was significant ($F(2,58) = 12.39$, $p < 0.001$), as was the effect of Trial Bin ($F(3,87) = 5.09$, $p < 0.003$). The interaction of the two effects was not significant ($F < 1$). The effect of trial bin was best described by a linear trend ($F(1,29) = 14.65$, $p < 0.001$), which accounted for more than 90% of the variance due to trial bin. Similarly, the RT difference was greater in Block 1 than in Block 6, for both the old and new locations ($F(1,58) = 23.67$, $p < 0.001$), which accounted for more than 95% of the variance due to Block.

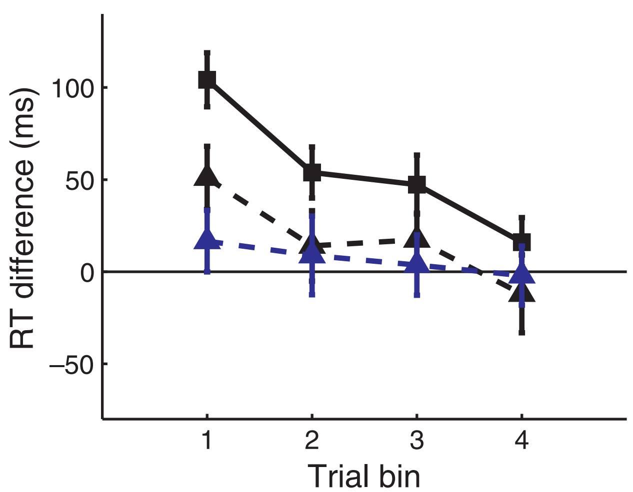

Figure 10. Difference in RT as a function of trial bin (see text for information on the binning of trial pairs). Black squares with solid line: Block 1; black triangles with dashed line: Block 6 old locations; blue triangles with dashed line: Block 6 new locations.

To further establish the presence of transfer effects, the RT difference due to distractors from the current study was compared with the data from Experiment 2, where practice effects did not transfer to new locations. We used a mixed factorial ANOVA that examined the between-subjects factor of Experiment (2 vs. 4) and the within-subjects factors of Session Half and Block. For this analysis, only the data from new distractor locations in Blocks 6–10 of Experiment 4 were used, to more closely match the conditions of Experiment 2. The effect of Experiment was significant ($F(1,58) = 21.59$, $p < 0.001$); this is expected given the large overall difference in the distractor effect. The main effect of Session Half was also significant ($F(1,58) = 5.48$, $p < 0.03$); there was also a trend toward a significant effect of Block ($F(4,232) = 2.35$, $p = 0.055$). There was also a trend for the interaction between Session Half and Experiment ($F(1,58) = 3.88$, $p = 0.054$), reflecting the large difference in the distractor effect after the introduction of the new distractor conditions in Experiment 4. This was confirmed by an independent-samples *t*-test examining the difference in distractor effect between Blocks 1 and 6 across the two experiments ($t(58) = 3.07$, $p < 0.005$). No other interactions were significant ($F_{\text{Exp} \times \text{Block}} < 1$; $F_{\text{Half} \times \text{Block}}(4,232) = 1.75$, $p = 0.14$; $F_{\text{Exp} \times \text{Half} \times \text{Block}}(4,232) = 1.77$, $p = 0.14$).

Discussion

This experiment replicated the central finding of the studies thus far, that repeated exposure to distracting stimuli reduces the impact of distraction on response times. In contrast to previous studies, though, this experiment showed that these practice effects transferred to new distractor locations. This difference appears to be the result of increased heterogeneity of the distractor items. Furthermore, the RT difference took longer to approach zero than in previous studies, suggesting that more heterogeneous distractors require more learning to achieve fully effective selectivity, but that this results in more generalized distractor filtering.

Two other methodological changes in this experiment might have contributed to these results. First, subjects here participated in a full session of the task in which distractors did not appear before seeing any distractors. It is possible that greater familiarity with the task led subjects to be more susceptible to attentional capture when distractor items were introduced in the second session. Though this would have little effect on the transfer of practice to new locations, it could have been a factor in the slower decline in the distractor effect. This would be consistent with the Attentional Load theory put forth by Lavie and colleagues (see Lavie, 2005, for a review), which states that tasks that place less of a demand on attentional resources are more open to disruption by distracting stimuli.

This possibility is unlikely for two reasons. First, the accuracy and RT at the end of Session 1 (and for all of Session 2) were within the range seen for all experiments thus far, as discussed previously. Second, the level of distraction at the beginning of the second session was also in line with the initial distraction effect for the other experiments (see Figure 11 and General discussion section), indicating that subjects were not more distracted in this experiment, but rather that they simply took longer to filter out the highly heterogeneous distracting stimuli.

The second methodological change in this experiment was that new distractor locations were introduced while distractors were still appearing in the old locations, rather than switching to new locations (cf. Experiment 2). It is possible that the continued appearance of distractors in their old locations made it easier to generalize to the new

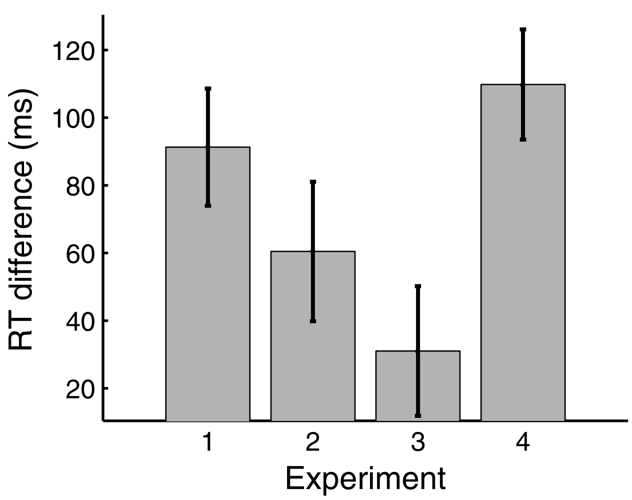

Figure 11. RT difference between Distractor and No Distractor trials for the first ten trial pairs of each experiment.

locations. Such a result seems unlikely, as the presence of new locations within the same block as old locations should be just as unexpected as the appearance of new locations on their own, as in the previous experiments. Furthermore, the data are inconsistent with the idea that simultaneously presenting old and new locations led to improved generalization. Such a generalization would take at least two distractor-present trials to develop (a new location trial to introduce the new condition, and an old location trial to establish the presence of the old condition) and in fact it would probably require several more trials. If this were the case, one would expect that distractors appearing in new locations would initially have a larger impact on RT. As can be seen in Figure 10, though, the new location trials (blue line) had a lower initial distractor effect than the old location trials (dashed black line) in Block 6; both old and new locations in Block 6 produced less distraction than that observed in Block 1. This held even for the first correct trial in each location condition, with the distraction effect being smaller in the new locations (22.3 ms) than the old locations (85.7 ms; a paired t-test failed to show any difference between these conditions however, $t(28) = 1.08$). Thus, subjects were able to efficiently filter out distractors appearing in new locations from the first such trial.

General discussion

In the experiments described here, subjects were asked to perform an attentionally demanding visual discrimination task. On some trials, a task-irrelevant, abrupt-onset distractor item would appear prior to the main task array. Initially, the appearance of these distractor items led to a reduction in the accuracy of responses and an increase in response time. After repeated exposure, the distractor-induced RT slowing declined (Experiment 1). When the appearance and location of the distractors were highly predictable, these practice effects were disrupted by the appearance of distractors in new locations (Experiment 2) and of new distractor items (Experiment 3). When the appearance of the distractors was more heterogeneous, subjects required more exposure to the task in order to overcome their distracting effect; however, these practice effects transferred more readily to new distractor conditions (Experiment 4). Thus, practice with a heterogeneous distractor set led to more generalized improvement in the selection efficiency.

The conclusion that subjects required more practice to overcome distraction with heterogeneous distractors is based on the fact that in Experiments 1–3, only the first block and the block immediately following a distractor switch (i.e., Blocks 1 and 6) exhibited an RT difference that was significantly greater than zero. In contrast, the RT difference in Experiment 4 lasted for several blocks. This difference might be due in part to a difference in the magnitude of the initial distraction effect. Figure 11 shows the RT difference for the first ten trial pairs in each of the above experiments. This plot does show some variability in the initial distraction level between the different experiments. A between-subjects ANOVA showed that Experiment had a significant effect on initial RT difference ($F(3,111) = 3.45$, $p < 0.05$). However, post-hoc analyses using Tukey's HSD only showed a significant difference between Experiments 3 and 4 ($p < 0.02$); no other pairs were significantly different (all ps > 0.17). As discussed earlier, the subjects in Experiment 3 were very efficient in overcoming the initial distraction effect. This suggests that the slower reduction in RT cost was not entirely due to the magnitude of the RT difference being greater at the beginning of the experiment.

The pattern of results observed here, that increased distractor variability leads to more effective transfer across distractor conditions, is not consistent with any of the explanations discussed previously. The fact that subjects were able to suppress distractors from multiple classes of objects, all of which were unique images, and transfer this suppression to new locations, strongly argues against sensory adaptation. The lack of transfer in the distractor item transfer condition (Experiment 3) and the transfer across locations in Experiment 4 rule out the idea of suppression of specific locations. The quick reduction in the distractor effect for disk/face distractors (Experiments 1–3) and the slower reduction when there are many classes of distractor item tend to support an object-based suppression mechanism. However, the difference in the levels of transfer to new locations in the different experiments seems to contradict this: more efficient suppression of distractors should in principle produce more effective transfer. This suggests that a different mechanism may be responsible for the reductions in the distractor effect.

On possible explanation for the observed results is habituation of the orienting response (OR), a process that has been extensively studied (Sokolov, 1963; see Bradley, 2009, for a review). The OR consists of various physiological and neural responses to the presence of emotional, salient, or novel stimuli. These responses are typically considered to be indicators that attention has been directed to, or captured by, the stimuli. The OR diminishes quickly with repetition of the stimuli but is reinstated with the presentation of another novel stimulus. This paradigm has typically been used to assess the significance of feature differences between two or more similar objects, similar to adaptation patterns in fMRI (as reviewed by Krekelberg et al., 2006).

Habituation of the OR may account for the reduced distractor effect observed in Experiments 1 and 2 and, to a lesser extent, in Experiment 3. In these experiments, a single stimulus (the red and green disks) is presented repeatedly across multiple trials, in relatively stable locations. These circumstances would promote rapid reduction in the OR. If subjects are no longer orienting to the distractor, it should cease to cause a slowing in the response to the main task. The uniformity of the disks then leads to reinstatement of OR following a location change. Elliott and Cowan (2001) explained a reduction in distraction during a cross-modal Stroop paradigm as due in part to a similar habituation of the OR to auditory distractors.

With the introduction of the more variable distractor sets in Experiments 3 and 4, the current results become less amenable to explanation by appealing to a habituated OR. As discussed previously, Gati and Ben-Shakhar (1990; as well as Bradley, 2009) present evidence that habituation can be based on repetition of specific features across multiple images, rather than just repetition of a single stimulus. However, the faces used in Experiment 3 were suboptimal stimuli for habituation to specific features, both in terms of presentation and set heterogeneity. This is doubly true for the distractors used in Experiment 4. One could argue that the increased time required to reduce the distraction effect in Experiment 4 represents gradual habituation to multiple feature sets across the full range of stimuli. However, the use of many different stimulus categories makes the number of features that would need to be habituated so large as to be implausible, or so general as to be perceptually untenable (e.g., habituation to curved or straight edges, etc.). This suggests that the observed reductions in the distractor effect, and their transfer to new spatial locations, are the result of improvements in the allocation of attention, rather than habituation of an orienting response. Future studies can further clarify this point by testing transfer of attentional learning effects to previously unseen distractor categories.

That subjects took longer to develop efficient selectivity with heterogeneous distractors suggests a system for distractor suppression that scales with the amount of information available. Where distractors are highly predictable and stable in their appearance, filtering mechanisms can be very narrow and specific to the particular stimuli to be ignored. Such a system would require inhibition of relatively few cortical regions corresponding to either object representation or spatial location, and so should be quickly reinforced over time, reaching peak efficiency with only a few repetitions. Additionally, the cognitive demands of such a filtering mechanism would be relatively low, making it an effective strategy for performing a task such as the one used here.

Where the distractors are highly variable and unpredictable, a stimulus-specific filtering mechanism would prove ineffective. Instead, a more general filtering mechanism would be required, one that could be applied in many different conditions. Because the first few distractor exposures would be similar to a predictable distractor condition (with limited sampling of the distractor set, the level of heterogeneity in the distractor set would be indeterminate), the filtering system would be established in a bottom-up fashion, updating as more information became available. Such a system would be slow to reach peak efficiency for two reasons. First, as new distractor items continued to appear, they would be outside the bounds of the system currently in place, and so would not be effectively filtered. This would initially yield only minimal reduction in the distractor effect until a sufficiently broad filter was established. Second, a general filtering mechanism would still require enough constraints that it did not impair perception of task-relevant information (e.g., the array of red and green dots). Thus, inhibiting all sensory input in a given modality could not be used. It follows then that a filtering mechanism that would work within the constraints of the current task would involve inhibitory interactions among a variety of visual processing areas, such as retinotopically mapped early visual cortex and object-selective inferior temporal cortex. Because these inhibitory connections would cover many functionally specialized regions, making the distractor filtering system more complex overall, any single distractor item would necessarily provide less information for the execution of the filtering system, compared to the case with highly predictable distractor items. Thus, less variable distractors would be overcome more quickly because the system for filtering out the distractor information would be more specific, and thus implemented more quickly. In contrast, a more general filtering mechanism would take longer to be established.

Such a mechanism, in which more generalized training leads to greater transfer to new distractor conditions, is supported by work in both verbal memory and motor learning, as reviewed by Schmidt and Bjork (1992). Here, the authors describe several studies investigating many different tasks showing that conditions that promote rapid improvement during practice do not lead to optimal post-training performance. Rather, training conditions that typically lead to worse performance during practice produce more complete learning and improved performance in a wider variety of post-training scenarios. The authors argue that this is because the more difficult and varied practice conditions lead to deeper and more complete information processing, necessarily leading to better performance. The present findings, and proposed models, are consistent with this idea, drawing an important parallel between attentional learning and skill training in general. This has important implications for future studies of attentional processes.

The results discussed here diverge from previous examinations of improvements in attentionally demanding tasks with repeated exposure. The experiments of Schneider and Shiffrin (1977) showed improvements in target detection when the set of target items was repeated over many trials; perturbations in this set eliminated all practice effects. The contextual cueing paradigm (Chun & Jiang, 1998, 1999) leads to improved task performance because the distractors actually guide attention, rather than being filtered (though see Makovski et al., 2008, for an example of learning restricted to the target items). Brown and Fera (1994) showed a reduction in the flanker effect with practice, but only under limited circumstances and with distractors that carried task-relevant information (the flankers were predictive of the main target); these effects also seemed to be highly dependent on the particular parameters and timing of task, suggesting limited possibility for transfer. These studies are examples of the general trend in tasks that examine practice effects and attention: improvements are tied to the particular task-relevant stimuli and generally involve the development of automatic associations between stimuli and the correct response.

A recent study by Dixon, Ruppel, Pratt, and De Rosa (2009) also examined the issue of improvement in the allocation of attention and filtering of distracting information. In their study, subjects were asked to identify which of a pair of objects belonged to a previously defined target set based on color and shape. During priming blocks, only color information was useful for identifying the target item, with the shape of the objects being non-informative. During the probe blocks, however, subjects needed to use the shape of the objects to identify the targets, and color became non-informative. Subjects were slower to respond to the shapes in the probe blocks after having ignored them in the prime blocks; this effect was largest when the shapes were previously ignored in two separate color contexts, indicating generalization of the filtering process. The findings of Dixon et al. are mirrored by the current results, which examine improved filtering not for features of attended objects but for completely separate distractor items. Of particular importance is the more generalized learning effect produced by a more variable training condition found in both studies, again corresponding to the processes described in other skill training by Schmidt and Bjork (1992).

As discussed earlier, there is some evidence for improvements in the performance of working memory tasks with practice, though it is not consistent. Where such evidence does exist (Jaeggi et al., 2008; Klingberg et al., 2002; Olesen et al., 2004), it seems to suggest that effective training on a single working memory task leads to general improvements in working memory capacity and overall cognitive ability. Because the current studies only investigated transfer across different distractor conditions, it remains to be seen whether improvements in distractor filtering lead to general improvements in selective attention in different task domains. As such, the relationship between the current results and improvements in working memory performance remains uncertain.

The improvements in distractor filtering observed here raise the possibility that perceptual learning may have been a factor. Studies of perceptual learning involve extensive practice with identification and discrimination of subtly different visual stimuli (for reviews, see Ahissar & Hochstein, 2004; Fine & Jacobs, 2002; Petrov, Dosher, & Lu, 2005). In these studies, the extent to which learning transfers to new conditions is highly dependent on the stimulus and task parameters. In their Reverse Hierarchy model, Ahissar and Hochstein (2004) have suggested that transfer of training is dependent on the perceptual processing level at which the learning took place. According to this account, greater transfer would be expected where relatively coarse discriminations can be accomplished by later processing stages (e.g., discrimination between objects, relying on IT cortex). However, more difficult discrimination requiring more fine-tuned analysis would produce less transfer because it relied on changes in a more narrowly responsive set of neurons (e.g., a single orientation column in area V1). Alternately, Dosher and Lu (1999, 2007; Petrov et al., 2005) argue that learning is the result of fine-tuning of perceptual templates, resulting in improved filtering of external noise. They further argue that such fine-tuning only happens under difficult discrimination conditions (i.e., with confusable targets in the presence of distracting information), and that because it manifests as a change in the strength of connections between perceptual and response processing stages, it is specific to the learned stimulus and the context in which it was viewed.

At first glance, the present data seem to contradict these models, given the greater transfer observed in Experiment 4 compared to Experiments 1–3. However, several aspects of the current paradigm make perceptual learning models a poor fit for these results. Though the distractors, being complex objects, would have been processed in object responsive regions of visual cortex (where the Reverse Hierarchy model states that generalized learning is more likely), there was no need for improved discrimination of these items. Their locations were predictable, and none of them resembled the target array, making the improved discrimination afforded by perceptual learning unnecessary for any improved filtering of the items. Furthermore, the distractors were spatially distinct from the dot array and several degrees into the periphery, reducing the degree to which they could visually interfere with the array. Lu, Lesmes, and Dosher (2002) showed that endogenous attention improved perceptual template mapping in much the same way as perceptual learning, by reducing the effects of external sensory noise. They also showed that this reduction in external noise is confined to the target location and seems to have little impact on adjacent locations when the target location is known in advance (as was the case here). This suggests that perceptual learning accounts are not adequate to explain the present results.

Though the proposed distractor filtering mechanism that develops based on information about the distractor set addresses how improvements in performance arise, it does not address the nature of those improvements. How does practice change the way in which the distractors are processed so that they no longer slow responses? The distracting items used here were all abrupt onsets, a class of stimuli that have been shown to effectively capture visuospatial attention (see Yantis & Jonides, 1996); additionally, the distractors share certain features with the task array (the red and green disk distractors have the same colors as the array; both distractors and the task array appear abruptly on the display) that could facilitate contingent capture based on subjects' attention priority sets (see, e.g., Folk & Remington, 1998). Because the distracting items successfully capture attention, they consume limited attentional resources, which are necessary for task performance, producing slower response times. Previous work on attentional capture, though, does not indicate that subjects become less susceptible to capture over time. This suggests that the reduction in the distractor effect may not be the result of a reduction in attentional capture.

A key difference between this paradigm and others is that most other experiments involving attentional capture show decrements to performance in situations of spatial uncertainty. For example, capture in a visual search task usually involves a stimulus that draws attention to its spatial location. In situations when this stimulus is not the target item, subjects must then continue searching for the target in the rest of the display. The capture of attention in this case has produced a delay in the response (see also Folk et al., 2002, for an example of capture in a situation of temporal uncertainty). In the current task, there is neither spatial nor temporal uncertainty. The task array always appears in the same location and at a very regular interval from one trial to the next. This certainty could allow subjects to adapt to the presence of the distracting items and eventually avoid being captured by them. If temporal and spatial regularities are important factors in overcoming attentional capture, then it follows that more regular distractors would be easier to overcome, as was the case here. In contrast, distractors that had a more variable appearance might require a more general filtering mechanism to avoid capture by task-irrelevant stimuli, accounting for the greater transfer of practice effects in those cases.

Alternately, it is possible that subjects are not overcoming the capture of attention, but rather the tendency to then process the identity and behavioral relevance of the distractors. In this case, the sensory information about the distractors might still be processed, but that information would no longer interfere with the cognitive operations of the main task. Such an improvement would also be tied to the level of variability in the distractor items, resulting from either the rapid adaptation of an information filtering mechanism that was specific to the observed distractors, or a more general one that could be applied to previously unobserved conditions.

These two explanations of the observed improvement in performance lead to two different predictions about how activity in the brain changes along with practice. If the distractors no longer capture subjects' visuospatial attention, then one would expect to see reduced activity in the early visual areas associated with the appearance of the distractors. In contrast, if practice is leading to reduced cognitive interference, then one should observe changes in the activity of more frontal regions, while the response in early visual areas remained stable. Such predictions could be addressed directly by adapting this paradigm for functional neuroimaging.

Selecting task-relevant and ignoring task-irrelevant information is a critical aspect of successful goal-directed behavior. The experiments described here addressed these issues by examining distraction effects during a visual discrimination task. The data revealed that repeated exposure to distracting stimuli led to an improved ability to filter out the distracting information. Furthermore, exposure to a more variable set of distractors produced more effective filtering that could be applied to previously unseen locations. These studies show that selective attention improves with training, and that these improvements occur at the cognitive, rather than the sensory, level. This relates improvements in attention to previous studies of skill training in general, which has important implications for the future study of attention and information selection.

Acknowledgments

This research was supported by the National Institutes of Health Grant R01-DA13165 to SY. The authors would like to thank Amy Shelton, Howard Egeth, and Charles Connor for helpful suggestions, and Jared Abrams and Heather Rosen for help with data collection.

Image sources: Grayscale faces, buildings, objects, and animals taken from Hemera Photo-Objects 50,000, vol. II (Dundas Software, 2000). Color faces taken from the faces94 set provided by Dr. Libor Spacek as part of the Computer Vision Science Research Project (http://cswww.essex.ac.uk/mv/allfaces/index.html). Color objects taken from The Object Databank provided by Dr. Michael Tarr (http://titan.cog.brown.edu:8080/TarrLab/stimuli/objects/). Fractals taken from stimulus set used by Dr. Joseph Sala and colleagues (Sala & Courtney, 2007).

Commercial relationships: none.
Corresponding author: Todd A. Kelley.
Email: t.kelley@ucl.ac.uk.
Address: Research Department of Cognitive Perceptual and Brain Sciences, Gower Street, London, United Kingdom.

References

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

Awh, E., Matsukura, M., & Serences, J. T. (2003). Top-down control over biased competition during covert spatial orienting. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 52–63. [PubMed]

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology, 46,* 1–11. [PubMed]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Broadbent, D. E., & Broadbent, M. H. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42,* 105–113. [PubMed]

Brown, P., & Fera, P. (1994). Turning selective attention failure into selective attention success. *Canadian Journal of Experimental Psychology, 48,* 25–57.

Chen, D., Eng, H. Y., & Jiang, Y. (2006). Visual working memory for trained and novel polygons. *Visual Cognition, 14,* 37–54.

Chua, K.-P., & Chun, M. M. (2003). Implicit scene learning is viewpoint-dependent. *Perception & Psychophysics, 65,* 72–80. [PubMed] [Article]

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Science, 4,* 170–178. [PubMed]

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. [PubMed]

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10,* 360–365.

Dixon, M. L., Ruppel, J., Pratt, J., & De Rosa, E. (2009). Learning to ignore: Acquisition of sustained attentional suppression. *Psychonomic Bulletin and Review, 16,* 418–423. [PubMed]

Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research, 39,* 3197–3221. [PubMed]

Dosher, B. A., & Lu, Z.-L. (2007). The functional form of performance improvements in perceptual learning. *Psychological Science, 18,* 531–539. [PubMed]

Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractor in a cross-modal color-word interference task. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 654–667. [PubMed]

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception & Psychophysics, 16,* 143–149.

Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning tasks: A review. *Journal of Vision, 2*(2):5, 190–203, http://journalofvision.org/2/2/5/, doi:10.1167/2.2.5. [PubMed] [Article]

Folk, C., Leber, A., & Egeth, H. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics, 64,* 741–753. [PubMed] [Article]

Folk, C., Remington, R., & Johnston, J. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance, 18,* 1030–1044. [PubMed]

Folk, C., Remington, R., & Wright, J. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception & Performance, 20,* 317–329. [PubMed]

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 847–858. [PubMed]

Furl, N., van Rijsbergen, N. J., Teves, A., & Dolan, R. J. (2007). Face adaptation aftereffects reveal anterior medial temporal cortex role in high level category representation. *Neuroimage, 37,* 300–310. [PubMed]

Gati, I., & Ben-Shakhar, G. (1990). Novelty and significance in orientation and habituation: A feature-matching approach. *Journal of Experimental Psychology: General, 119,* 251–263. [PubMed]

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10,* 14–23. [PubMed]

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 Science, 105,* 6829–6833. [PubMed] [Article]

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

Krekelberg, B., Boynton, G. M., & van Wezel, R. J. (2006). Adaptation: From single cells to BOLD signals. *Trends in Neuroscience, 29,* 250–256. [PubMed]

Lamy, D., Leber, A., & Egeth, H. (2004). Effects of task relevance and stimulus-driven salience in feature-search mode. *Journal of Experimental Psychology: Human Perception & Performance, 30,* 1019–1031. [PubMed]

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Science, 9,* 75–82. [PubMed] [Article]

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22,* 1–35.

Lu, Z.-L., Lesmes, L. A., & Dosher, B. A. (2002). Spatial attention excludes external noise at the target location. *Journal of Vision, 2*(4):4, 312–323, http://journalofvision.org/2/4/4/, doi:10.1167/2.4.4. [PubMed] [Article]

Makovski, T., Vázquez, G. A., & Jiang, Y. V. (2008). Visual learning in multiple-object tracking. *PLoS ONE, 3,* e2228. [PubMed] [Article]

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

Olson, I. R., & Chun, M. M. (2001). Temporal contextual cueing of visual attention. *Journal of Experimental Psychology: Learning, Memory & Cognition, 27,* 1299–1313. [PubMed]

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9,* 273–302.

Olson, I. R., Chun, M. M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early, temporally late stages of visual processing. *Brain, 124,* 1417–1425. [PubMed]

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. [PubMed]

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review, 112,* 715–743. [PubMed]

Remington, R., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics, 51,* 279–290. [PubMed]

Sala, J. B., & Courtney, S. M. (2007). Binding of what and where during working memory maintenance. *Cortex, 43,* 5–21. [PubMed]

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207–218.

Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84,* 1–66.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science, 16,* 114–122. [PubMed]

Serences, J. T., Yantis, S., Culberson, A., & Awh, E. (2004). Preparatory activity in visual cortex indexes distractor suppression during covert spatial orienting.

Journal of Neurophysiology, 92, 3538–3545. [PubMed] [Article]

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127–190.

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. New York: Pergamon Press.

Tang, Y.-Y., & Posner, M. I. (2009). Attention training and attention state training. *Trends in Cognitive Sciences, 13,* 222–227. [PubMed]

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428,* 557–561. [PubMed]

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception & Performance, 25,* 661–676. [PubMed]

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance, 10,* 601–621. [PubMed]

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary vs. automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance, 16,* 121–134. [PubMed]

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt visual onsets: New perceptual objects or visual masking? *Journal of Experimental Psychology: Human Perception & Performance, 22,* 1505–1513. [PubMed]