Chapter 28

EFFECTS OF CONTEXT AND INSTRUCTION ON THE GUIDANCE OF EYE MOVEMENTS DURING A CONJUNCTIVE VISUAL SEARCH TASK

JIYE SHEN, AVA ELAHI-PANAH, and EYAL M. REINGOLD

University of Toronto, Mississauga, Canada

Contents

Abstract 596
1. Subset-selective processing and distractor-ratio effect 597
2. Bottom-up and top-down processing 598
3. Method 601
   3.1. Participants 601
   3.2. Apparatus 601
   3.3. Stimuli and design 601
   3.4. Procedure 603
4. Results and discussion 603
   4.1. Response time 604
   4.2. Number of fixations per trial 606
   4.3. Initial saccadic latency 606
   4.4. Fixation duration 607
   4.5. Saccadic selectivity 608
5. General discussion 609
Acknowledgements 611
References 611

Eye Movements: A Window on Mind and Brain
Copyright © 2007 by Elsevier Ltd. All rights reserved.
Abstract

The current study explored the effects of contextual and instructional manipulations on visual search behavior in a conjunctive search task with a distractor-ratio manipulation. Participants’ eye movements were monitored when they performed the task. Results from the present investigation demonstrated that block context modulates the temporal dynamics of visual search by shortening fixation durations, reducing the number of fixations, enhancing task-relevant subset selection, and as a result, yielding faster manual responses in those trials that were congruent with the block context than in incongruent trials. The current study also suggests that the presence of top-down influences associated with an instructional manipulation operates over and above any influences of context.
The current study explored the effects of contextual and instructional manipulations on visual search behavior in a conjunctive search task with a distractor-ratio manipulation. In a typical conjunctive visual search task, each trial contains an equal number of distractors of each type; the total number of items within a search display (display size) is manipulated and search efficiency is examined by measuring the change in response time and/or error rate as a function of display size (see Treisman, 1988; Wolfe, 1998 for a review). However, several recent studies have shown that visual search performance is sensitive to the ratio between the types of distractors, even when the total number of items in a display remains constant (e.g., Bacon & Egeth, 1997; Egeth, Virzi, & Garbart, 1984; Kaptein, Theeuwes, & van der Heijden, 1995; Poisson & Wilkinson, 1992; Shen, Reingold, & Pomplun, 2000, 2003; Sobel & Cave, 2002; Zohary, & Hochstein, 1989).

1. Subset-selective processing and distractor-ratio effect

The original feature-integration theory of visual search (Treisman & Gelade, 1980; Treisman, 1988) proposes the existence of preattentive feature maps, one for each stimulus dimension (such as color, shape, orientation, etc.). Information from parallel preattentive processes could only mediate performance in a visual search task if the target was defined by the presence of a unique feature (i.e., feature search), such as searching for a green \(X\) among red and blue \(X\)s. However, if the target is defined by a specific combination of features (i.e., conjunction search), such as searching for a green \(X\) among red \(X\)s and green \(O\)s, attention is necessary to locally combine the information from the corresponding feature maps. As a result, participants have to inspect the search display in a serial item-by-item fashion until target detection or exhaustive search.

In contrast to the arguments by the original feature-integration theory, subsequent studies have demonstrated that participants could selectively limit their search to only a subset of distractors to improve search efficiency (e.g., Carter, 1982; Egeth et al., 1984; Friedman-Hill & Wolfe, 1995; Kaptein et al., 1995). For example, Egeth et al. (1984) had participants search for a red \(O\) (target) among red \(N\)s (same-color distractors) and black \(O\)s (same-shape distractors). They varied the number of distractors of one type (i.e., same-color or same-shape) while keeping the number of distractors of the other type constant. Participants were instructed to try to restrict their search to the subset of items that were kept constant in number. Egeth et al. found that search times were independent of the number of items in the uninstructed subset, regardless of whether participants were instructed to attend to the color subset or the shape subset. Similarly, studies that have examined patterns of eye movements during visual search have provided strong evidence of bias in the distribution of saccadic endpoints across different types of distractors. Stimulus dimensions such as color, shape, contrast polarity, and size have been shown to guide the search process (e.g., Bichot & Schall, 1998; Findlay, 1997; Findlay & Gilchrist, 1998; Hooge & Erkelens, 1999; Motter & Belky, 1998; Pomplun, Reingold, & Shen, 2001, 2003; Scialfa & Joffe, 1998; Shen et al., 2000; Shen, Reingold, Pomplun, & D. E. Williams, 2003; D. E. Williams & Reingold, 2001).
Evidence for subset-selective processing also comes from studies using the *distractor-ratio manipulation*. In a color × orientation conjunction search task, Zohary and Hochstein (1989) asked participants to decide whether a red horizontal bar was present among an array of red vertical (same-color distractors) and green horizontal (same-orientation distractors) bars. The search display was presented very briefly (50 ms) and then, after a variable interval (stimulus onset asynchrony, SOA), masked. One critical manipulation in this study was the ratio between the two types of distractors (same-color vs same-orientation) presented in a given array. The number of same-color distractors ranged from 0 to 64, in increments of 4, while the total number of items was held constant at 64. Zohary and Hochstein found that the SOA required to reach a 70% correct response rate was a quadratic function of the number of distractors sharing color with the search target. Specifically, detection was relatively easy for displays with extreme distractor ratios (i.e., either the same-color or same-orientation distractors were rare) but relatively difficult for displays in which the two types of distractors were equally represented. In addition, these investigators found that the performance curve was not completely symmetrical (it took a shorter SOA for participants to reach a certain level of accuracy for displays with few same-color distractors than for displays with a comparable number of same-orientation distractors). The finding that visual-search efficiency in a conjunctive search task depends on the relative frequency of the two types of distractors is consistent with the notion of subset-selective processing and has been referred to as the *distractor-ratio effect* (Bacon & Egeth, 1997).

In addition to the accuracy across SOA measure (Zohary & Hochstein, 1989), the distractor-ratio effect has also been observed in studies measuring response times (Bacon & Egeth, 1997; Poisson & Wilkinson, 1992; Sobel & Cave, 2002). Shen et al. (2000) further examined participants’ patterns of eye movements, the spatial distribution of saccadic endpoints in particular, during the search process. They employed a color × shape conjunction search task and systematically manipulated the ratio between the same-color and same-shape distractors in a display. They found a quadratic change in search performance measures such as manual response time, number of fixations per trial, and initial saccadic latency as a function of distractor ratio. Search performance was worse when the ratio between the same-color and same-shape distractors approximated 1:1 and gradually improved as the ratio deviated from 1:1, with performance being best at extreme distractor ratios (i.e., very few distractors of one type). More importantly, Shen et al. demonstrated that when there were very few same-color distractors, participants’ saccadic endpoints were biased towards the color dimension whereas when there were very few same-shape distractors, saccades were biased towards the shape dimension. Results from that study suggest that in a distractor-ratio paradigm, participants take advantage of the display information and flexibly switch between different subsets of distractors on a trial-by-trial basis.

### 2. Bottom-up and top-down processing

Visual attention is currently thought to be controlled by two distinct mechanisms: one is the top-down or goal-directed control, in which the deployment of attention is determined...
by the observer’s knowledge or intentional state, and the other is bottom-up or stimulus-driven control, in which attention is driven by certain aspects of the stimulus, irrespective of the observer’s current goals or intent (Egeth & Yantis, 1997; Wolfe, Butcher, Lee, & Hyle, 2003; Yantis, 1998). Various attention and visual search studies have proposed the involvement of both the top-down and bottom-up processes in determining attentional allocation and visual search performance (e.g., Bacon & Egeth, 1994, 1997; Bravo & Nakayama, 1992; Cave & Wolfe, 1990; Hillstrom, 2000; Irwin, Colcombe, Kramer, & Hahn, 2000; Wolfe, 1994).

The finding that participants were able to select and search through the smaller subset of distractors, even when different levels of distractor ratios were randomly mixed in the same block of trials suggests that bottom-up processing driven by the saliency of display items (Koch & Ullman, 1985) plays a role in mediating the distractor-ratio effect (Wolfe, 1994). This has been supported by studies manipulating stimulus discriminability directly. For example, Shen et al. (2003) varied stimulus discriminability along the shape dimension in a color × shape conjunction task (X vs O was used in the high-discriminability condition whereas in the low-discriminability condition, the shapes were X vs K for half of the participants and O vs Q for the other half). These investigators found that, in the high-discriminability condition, participants searched through different subsets of distractors on the basis of distractor ratio (i.e., color subset for displays with very few same-color distractors and shape subset for displays with very few same-shape distractors). In marked contrast, in the low-discriminability condition, saccades were consistently biased towards the color dimension, irrespective of the distractor-ratio manipulation. This suggests that participants searched through the more informative, but not necessarily smaller, subset of distractors. Sobel and Cave (2002) similarly found that in a color × orientation conjunction task, when the difference between orientations was large, participants were able to search through the smaller subset of distractors whereas when orientation differences were small, participants tended to search through the subset of distractors sharing the target color.

The influence of top-down controls in the distractor-ratio effect has been examined in a recent study by Bacon and Egeth (1997), which provided participants with instructions designed to induce particular search strategies (see also Egeth et al., 1984; Kaptein et al., 1995; Sobel & Cave, 2002). With a color × orientation conjunction search task, these investigators manipulated the frequency of trials with very few same-color distractors and trials with very few same-orientation distractors. Participants were informed of which distractor type would be less frequent on most trials and were asked to restrict their search to a subset of distractors (e.g., attend to red items or attend to vertical items). To demonstrate the role of top-down processing in the distractor-ratio effect, these investigators contrasted two experimental conditions, one in which the instruction to the participants was congruent with the optimal search strategy for a given trial (i.e., top-down and bottom-up processes working in concert) with another condition in which the instruction was incongruent with the optimal search strategy (i.e., top-down and bottom-up processes working in opposition). Bacon and Egeth found that search performance was poorer when
the optimal search strategy and instruction mismatched than when they matched, and thus
demonstrated the role of top-down processing in the distractor-ratio paradigm.

The current study extended Bacon and Egeth (1997) by further investigating the influ-
ence of top-down controls on search performance in a distractor-ratio paradigm. With
a color × shape conjunction search task, search displays which contained either very
few same-color distractors (henceforth color-search displays) or very few same-shape
distractors (henceforth shape-search displays) were created. In each block of trials, the
frequency of these two types of search displays was manipulated, creating two types of
block context. Block context refers to the fact that, on average, the previously experienced
trials will have had few distractors with the relevant feature. In half of the blocks, 80%
of the trials were color-search displays and 20% of the trials were shape-search displays.

In these blocks, color-search displays were congruent with the block context (henceforth,
color-congruent trials) whereas the shape-search displays were incongruent with the block
context (henceforth, shape-incongruent trials). In the other half of the blocks, 80% of the
trials were shape-search displays and 20% of the trials were color-search displays. Thus,
in these blocks, color-search displays were incongruent with the block context (hence-
forth, color-incongruent trials) whereas the shape-search displays were congruent with
the block context (henceforth, shape-congruent trials). For both color-search displays and
shape-search displays, performance on the congruent trials and incongruent trials was
 contrasted. A finding of better search performance in the congruent trials than in the
incongruent trials (i.e., a congruency effect) would reflect the influence of block context
(e.g., Bichot & Schall, 1999; Chun & Jiang, 1999; Irwin, Colcombe, Kramer, & Hahn,
2000; Hillstrom, 2000; see Chun, 2000 for a review).

Furthermore, two groups of participants were included in this experiment. In one
condition, similar to Bacon and Egeth (1997), Egeth et al. (1984), and Kaptein et al.
(1995), participants were informed of the composition of trials and provided explicit
instructions regarding the optimal search strategies (i.e., attend to the color or attend
to the shape). For another group of participants, exactly the same search trials were
employed except that no explicit instructions concerning search strategies were provided.
A comparison across the two conditions permitted a separate evaluation of the influences
on search performance due to instruction vs the influences due to block context.

The second goal of the current study is to examine how the instructional and contextual
manipulations, if observed, influence the spatiotemporal dynamics of visual search behav-
ior. The above-mentioned studies that have examined the top-down controls (Bacon &
Egeth, 1997; Egeth et al., 1984; Kaptein et al., 1995; Sobel & Cave, 2002) focused on
global search performance measures such as response time and error rate only. Therefore,
it remains unknown exactly what aspects of search behavior are changed owing to the
top-down controls and whether there are qualitative and quantitative differences in search
behavior between the two forms (contextual vs instructional) of top-down controls. For
example, could the contextual or instructional congruency effect in manual response time
be due to less time spent, upon the presentation of the search display, to segment the
search display and decide upon which subset to search through? Could it be that the
availability of top-down information makes the rejection of distractor items easier? Or
could it be due to a higher bias in the selection of the relevant subset of distractors in the congruent trials, which leads to fewer fixations on the uninformative subset of distractors? To explore these possibilities, the current study also examined participants’ eye movements during the search process.

3. Method

3.1. Participants

Twenty-four undergraduate students, with normal or corrected-to-normal visual acuity and normal color vision, participated in three one-hour sessions. They were paid $30 for their participation. A between-subject design was adopted, with half of the participants performing in the instruction condition and the other half in the no-instruction condition. Although this design is less powerful than the within-subject design due to individual differences within each group, it permitted a separate evaluation of the influences on search performance due to instruction vs the influences due to block context without contamination.

3.2. Apparatus

The experiment was run in a lighted room and the luminance of the walls was approximately 30 cd/m². The eyetracker employed in the current study was the SR Research Ltd. EyeLink I system. This system has high spatial resolution (0.005°) and a sampling rate of 250 Hz (4-ms temporal resolution). Stimulus displays were presented on two monitors, one for the participant (a 19-inch Samsung SyncMaster 900P monitor with a refresh rate of 120 Hz and a screen resolution of 800 × 600 pixels) and the other for the experimenter. The experimenter monitor was used to give feedback in real-time about the participant’s computed gaze position. This allowed the experimenter to evaluate system accuracy and to initiate a recalibration if necessary. In general, the average error in the computation of gaze position was less than 0.5°. Participant made a response by pressing one of two buttons on a response box connected to the EyeLink I system.

3.3. Stimuli and design

Similar to Shen et al. (2000, 2003), display items were created by combining features on two stimulus dimensions: color (red vs green) and shape (X vs O). All display items were presented in a 15.5° × 15.5° field at a viewing distance of 91 cm. Each individual item subtended 0.8° both vertically and horizontally; the minimum distance between neighboring items was 2.0°. Participants were asked to search for the target item, a green X, among distractors that had either the same color (green Os, same-color distractors) or the same shape (red Xs, same-shape distractors) as the target. An equal number of target-present and target-absent trials were used. In each trial, the search display contained either
very few color-search displays (see Figure 1a for an example) or very few shape-search
displays (see Figure 1b for an example). For both types of search displays, the number
of items belonging to the smaller subset was 12 on average (8, 12, or 16) while the total
number of items within a display was fixed at 48. The CIE coordinates for the colors
were (.582, .350) for red and (.313, .545) for green.

One critical manipulation of the current study was the frequency of the color-search and
shape-search displays in a block of trials. In half of the blocks, 80% of the trials presented
color-search displays and 20% of the trials presented same-shape distractors, while the
opposite distribution of trial types occurred in the other half of blocks. The type of search
displays that occurred more frequently within a block of trials was labeled as “congruent”
and the infrequent type as “incongruent”. Color-congruent and shape-congruent blocks
were tested in alternate blocks. The current experiment also manipulated the instructions
given to the participants. In one condition (henceforth, instruction condition), participants
were provided additional instructions regarding the composition of search trials in a block
and the search strategies that worked for the congruent trials (e.g., in a color-congruent
block, subjects were given the following instructions: “In this block, most of the displays
contain fewer green O’s than red X’s. Therefore, you can maximize your search efficiency
by paying attention to the green items”). In the other condition (henceforth, no-instruction
condition), no such instructions were provided to the participants.

A four-factor factorial design was implemented in the current study, with target presence
(present vs absent), display type (color-search vs shape-search), trial congruency (congru-
ent vs incongruent) as within-subject factors and instructional manipulation (instruction
vs no instruction) as the between-subject factor. Participants performed a total of 2,160
trials across three individual sessions, which amounted to 432 congruent trials and 108

![Figure 1. Sample search displays used in the current study (target was a green X and the distractors were green O’s and red X’s). Red items shown in black; green items shown in white. (a) The panel illustrates a “color-search” display (target present). (b) The panel shows an example of “shape-search” display (target absent).](image-url)
incongruent trials for each combination of target presence and display type. At the beginning of each session, participants received two practice blocks of 16 trials.

3.4. Procedure

Participants were informed of the identities of the search target and distractor items before the experiment started. They were asked to look for the target item, and indicate whether it was in the display or not by pressing an appropriate button as quickly and accurately as possible.

A 9-point calibration procedure was performed at the beginning of the experiment, followed by a 9-point calibration accuracy test. Calibration was repeated if any point was in error by more than 1° or if the average error for all points was greater than 0.5°. Each trial started with a drift correction in the center of the computer screen and then press a start button to initiate a trial. The trial terminated if participants pressed one of the response buttons or if no response was made within 20 s. The time between display onset and the participant’s response was recorded as the response time.

4. Results and discussion

Because of an erroneous response, 4.9% (3.3% in target-absent trials and 6.5% in target-present trials) of trials in the instruction condition and 2.0% (0.8% in target-absent trials and 3.2% in target-present trials) of trials in the no-instruction condition were removed. In addition, those trials with a saccade or a blink overlapping the onset of the search display or with an excessively long or short response time (more than 3.0 standard deviations above or below the mean for each cell of the design) were excluded from further analysis. These exclusions accounted for 2.8 and 2.5% of all trials respectively in the instruction condition, and 1.9 and 2.9% of all trials respectively in the no-instruction condition.

The current study examined how the instructional and contextual manipulations influence the spatiotemporal dynamics of visual search behavior by analyzing response time and eye movement measures such as number of fixations per trial, latency to move, fixation duration, and saccadic bias. Both response time and fixation number are global search performance measures on how efficiently participants can determine target presence in an array. The initial saccadic latency and fixation duration provide fine-grained temporal information on the search process. These analyses were to extend previous investigation on the patterns of eye movements in visual search tasks (e.g., Binello, Mannan, & Ruddock, 1995; D. E. Williams, Reingold, Moscovitch, & Behrmann, 1997; Zelinsky & Sheinberg, 1997). All these measures were analyzed with a repeated-measures ANOVA, with target presence (2: target present vs target absent), display type (2: color-search vs shape-search), and trial congruency (2: congruent vs incongruent) as within-subject factors and instructional manipulation (2: instruction vs no instruction) as the between-subject factor. In addition, the bias in the distribution of saccadic endpoints (e.g., Findlay, 1997;

4.1. Response time

Table 1 shows the average response time as a function of target presence, display type, and trial congruency in both the instruction condition and the no-instruction condition. The repeated-measures ANOVA revealed a significant main effect of target presence, $F(1,22) = 107.63$, $Mse = 136863.50$, $p < 0.001$, indicating that response time was shorter in target-present trials than in target-absent trials. The instruction condition and the no-instruction condition did not differ in overall response time, $F < 1$. Consistent with our previous studies (Shen et al., 2000, 2003), search was more efficient for the color-search displays than for the shape-search displays, $F(1,22) = 20.48$, $Mse = 19292.14$, $p < 0.001$, and this difference was more pronounced in the target-absent trials than in the target-present trials, as indicated by a significant display type × target presence interaction, $F(1,22) = 23.60$, $Mse = 17851.35$, $p < 0.001$.

More importantly, the repeated-measures ANOVA revealed a significant effect of trial congruency, $F(1,22) = 33.86$, $Mse = 21973.47$, $p < 0.001$, indicating that response time was shorter in those trials that were either consistent with the instruction or the block.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Target-absent Trials</th>
<th></th>
<th>Target-present Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Color search</td>
<td>Shape search</td>
<td>Color search</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>IT</td>
<td>CT</td>
</tr>
<tr>
<td><strong>Response time</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instruction</td>
<td>1197</td>
<td>1416</td>
<td>1369</td>
</tr>
<tr>
<td>No instruction</td>
<td>1282</td>
<td>1346</td>
<td>1410</td>
</tr>
<tr>
<td><strong>Number of fixations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instruction</td>
<td>4.28</td>
<td>5.02</td>
<td>4.83</td>
</tr>
<tr>
<td>No instruction</td>
<td>4.68</td>
<td>4.89</td>
<td>5.14</td>
</tr>
<tr>
<td><strong>Initial saccadic latency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instruction</td>
<td>224</td>
<td>225</td>
<td>232</td>
</tr>
<tr>
<td>No instruction</td>
<td>224</td>
<td>225</td>
<td>229</td>
</tr>
<tr>
<td><strong>Fixation duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instruction</td>
<td>196</td>
<td>205</td>
<td>202</td>
</tr>
<tr>
<td>No instruction</td>
<td>195</td>
<td>199</td>
<td>199</td>
</tr>
</tbody>
</table>

J. Shen et al.
Figure 2. Congruency effect (the difference between the congruent trials and in congruent trials) for response time (a), number of fixations per trial (b), initial saccadic latency (c), and fixation duration (d) as a function of target presence, display type, and instructional manipulation.

context (congruent trials) than in those incongruent trials. Figure 2a plots the amount of congruency effect (i.e., the value of the congruent trials subtracted from that of the incongruent trials) as a function of target presence and display type, in both the instruction condition and the no-instruction condition. The overall congruency effect was greater in target-absent trials than in target-present trials, as indicated by a significant target presence × trial congruency interaction, $F(1, 22) = 11.04$, $Mse = 11423.61, p < 0.001$. The figure also shows that in target-present trials the instruction condition produced the same amount of congruency effect as did the no-instruction condition. However, in target-absent trials, the congruency effect was much stronger in the instruction condition than in the no-instruction condition. These findings were confirmed by a significant interaction of instructional manipulation × target presence × trial congruency, $F(1, 22) = 4.73$, $Mse = 11423.61, p < 0.05$. In addition, the congruency effect for the color-search displays was slightly larger in the shape-search displays, $F(1, 22) = 1.95$, $Mse = 8328.31, p = 0.176$. 
4.2. Number of fixations per trial

Table 1 shows that the instruction and no-instruction conditions did not differ in the average number of fixations per trial, $F < 1$. Similar to the response-time data, more fixations were made in the target-absent trials than in the target-present trials, $F(1, 22) = 153.49$, $Mse = 1.91, p < 0.001$. In addition, more fixations were made in the shape-search displays than in the color-search displays, $F(1, 22) = 16.66$, $Mse = 0.23, p < 0.001$; the difference between the two types of search displays was more pronounced in target-absent trials than in target-present trials, as indicated by a significant display type $\times$ target presence interaction, $F(1, 22) = 29.75$, $Mse = 0.21, p < 0.001$.

Figure 2(b) plots the congruency effect for fixation number as a function of target presence, display type, and instructional manipulation. The positive congruency values shown in the figure indicate that fewer fixations were made in the congruent trials than in the incongruent trials, as indicated by a significant main effect of trial congruency, $F(1, 22) = 32.81$, $Mse = 0.28, p < 0.001$. Overall, this congruency effect was stronger in target-absent trials than in target-present trials, with a significant interaction between trial congruency and target presence, $F(1, 22) = 8.73$, $Mse = 0.17, p < 0.01$. In addition, although the instruction condition produced the same amount of congruency effect as did the no-instruction condition in target-present trials, the congruency effect in target-absent trials was much stronger in the instruction condition than in the no-instruction condition. This was indicated by a significant interaction of instructional manipulation $\times$ target presence $\times$ trial congruency, $F(1, 22) = 4.22$, $Mse = 0.17, p < 0.05$. The figure also shows that in target-absent trials, the congruency effect was stronger for shape-search displays than for color-search displays, whereas there was no difference in the amount of congruency effect between the two types of search displays in target-present trials. This was confirmed by a significant target presence $\times$ display type $\times$ trial congruency interaction, $F(1, 22) = 6.16$, $Mse = 0.07, p < 0.05$.

Results from the current and previous parts reveal a high degree of similarity between response time and number of fixation data. This is not surprising given that both measures reflected global measures of search efficiency and that response time in a visual search task can be accounted for by overt shifts of attention (Findlay & Gilchrist, 1998; Scialfa & Joffe, 1998; D. E. Williams, et al., 1997; Zelinsky & Sheinberg, 1997).

4.3. Initial saccadic latency

Initial saccadic latency is defined as the interval between the onset of a search display and the detection of the first eye movement; this measure may be indicative of the ease of segmenting a search display and identifying an optimal search strategy. Initial saccadic latency was analyzed as a function of target presence, display type, and trial congruency (see Table 1). As in Shen et al. (2000, 2003), the initial saccadic latency was shorter in the color-search displays than in the shape-search displays, $F(1, 22) = 12.48$, $Mse = 95.51, p < 0.01$, and this effect was more pronounced in the target-absent trials than...
in the target-present trials, as indicated by a significant target presence × display type interaction, $F(1, 22) = 5.78$, $Mse = 29.26$, $p < 0.05$.

Figure 2(c) plots the amount of congruency effect for the initial saccadic latency. It is clear from the figure that the instructional and contextual manipulation did not have a substantial influence on the initial saccadic latency: There was neither a significant main effect for instruction manipulation, trial congruency, nor a significant interaction between instruction manipulation or trial congruency with other factors, all $F$s < 1.76, $ps > 0.199$. This suggests that the execution of first saccades is more influenced by the low-level display characteristics (color-search displays vs shape-search displays). Instead of remaining fixated in the center of the screen to gather initial target information or to check whether the current search display is consistent with optimal search strategy, participants tended to start the search right away (Zelinsky & Sheinberg, 1997). This suggests that top-down controls of visual attention modify the search behavior at a later stage.

4.4. Fixation duration

The repeated-measures ANOVA revealed that the instructional manipulation did not influence the overall fixation duration, $F < 1$. It is clear from Table 1 that average fixation duration was longer in the target-present trials than in the target-absent trials, $F(1, 22) = 10.04$, $Mse = 326.26$, $p < 0.001$. Consistent with earlier studies (e.g., Galpin & Underwood, in press; Pomplun, Sicheschmidt, Wagner, Clermont, Rickheit, & Ritter, 2001), the longer fixations in the target-present trials were linked to target detection. There was a significant main effect of display type, $F(1, 22) = 18.36$, $Mse = 44.99$, $p < 0.001$, indicating that fixation duration was longer in the shape-search displays than in the color-search displays. The longer fixation duration, together with more fixations made, may account for the relatively inefficient performance for the shape-search displays compared to the color-search displays.

The repeated-measures ANOVA also revealed a congruency effect for fixation duration, $F(1, 22) = 6.25$, $Mse = 40.73$, $p < 0.05$, with shorter fixation durations in the congruent trials than in the incongruent trials. There was also a significant instruction manipulation × target presence × display type × trial congruency interaction, $F(1, 22) = 8.13$, $Mse = 47.38$, $p < 0.05$. Figure 2(d) shows the amount of congruency effect for each cell of the design (target presence × display type × instructional manipulation). As can be seen from the figure, the fixation-duration congruency effect was consistently found in the target-absent trials, regardless of the display type or instructional manipulation. In target-present trials, however, a reliable congruency effect was observed only in the shape-search displays in the no-instruction condition. The figure also shows a reversed congruency effect for the color-search displays in the no-instruction condition. This unexpected finding was probably due to a surprisingly large reversal among three participants in that condition. The effects of congruency on fixation duration suggest that the availability of accurate top-down information makes the rejection of distractor items easier and thus speeds up the search process.
4.5. Saccadic selectivity

Was the distribution of saccadic endpoints influenced by the trial congruency manipulation? If the guidance of eye movements were solely driven by the salience of display items, the same level of saccadic selectivity should be observed across the congruent and incongruent trials, given that exactly the same search displays were used (i.e., with the same bottom-up activation). To examine this, bias in saccadic distribution for each type of search display was calculated – the baseline probability of fixation 25% (i.e., the number of items belonging to the smaller subset, 12 on average, out of a fixed display size of 48 items) was subtracted from the proportion of fixations on the same-color distractors for the color-search displays and from the proportion of fixations on the same-shape distractors for the shape-search displays. As pointed out by Zelinsky (1996; see also D. E. Williams & Reingold, 2001; Shen et al., 2000, 2003), results from target-absent trials can be interpreted more clearly than those from target-present trials where the presence of the target item may influence search behavior. As a result, only target-absent trials were included in the current analysis. Figure 3 plots saccadic bias as a function of display type and trial congruency in both the instruction condition and the no-instruction condition.

A repeated-measures ANOVA was conducted on saccadic bias, with display type (2: color-search vs shape-search) and trial congruency (2: congruent vs incongruent) as within-subject factors and instructional manipulation (2: instruction vs no instruction) as the between-subject factor. Overall, saccadic bias did not differ between the instruction condition and the no-instruction condition, $F(1, 22) = 2.22$, $Mse = 101.50$, $p = 0.15$. Consistent with the findings from Shen et al. (2000, 2003), saccadic bias was greater in

Figure 3. Congruency effect for saccadic bias as a function of display type and instructional manipulation (see text for the details on saccadic bias calculation).
the color-search displays, $F(1, 22) = 10.56$, $Mse = 193.32$, $p < 0.001$, implicating color
dominance in conjunction search tasks (see also Luria & Strauss, 1975; Motter & Belky,
1998; Pomplun et al., 2001, 2003; L. G. Williams, 1966; D. E. Williams & Reingold,
2001).

The repeated-measures ANOVA revealed a significant main effect of trial congruency,
$F(1, 22) = 25.96$, $Mse = 45.35$, $p < 0.001$, indicating stronger saccadic bias in the con-
gruent trials than in the incongruent trials. That is, participants were more likely to direct
saccades towards the same-color distractors for color-search displays and towards the
same-shape distractors for shape-search displays in those trials that were consistent with
the block context or instructional manipulation. This indicates that saccadic selectivity
was influenced not only by the bottom-up factors (salience of the display items) but also
by the top-down controls (instruction on search strategies or the contextual information).
In addition, Figure 3 also shows that the congruency effect was stronger in the instruction
condition (with a difference of 8.7% in the overall saccadic bias between the congruent
and incongruent trials, $t(11) = 5.03$, $p < 0.001$) than in the no-instruction condition (with
a difference of 1.6% in the overall saccadic bias between the congruent and incongruent
trials, $t(11) = 2.85$, $p < 0.05$). This was verified by a significant instructional manipula-
tion and trial congruency interaction, $F(1, 22) = 10.19$, $Mse = 45.35$, $p < 0.001$.

5. General discussion

The current study examined the top-down and bottom-up controls of visual attention in a
search task with distractor-ratio manipulation. Although exactly the same search displays
were used, search performance differed when the top-down information works in concert
or against the bottom-up information. Participants’ manual response time was shorter in
those trials that were congruent with the search instruction or block context than in the
incongruent trials. This suggests that visual guidance in a distractor-ratio experiment is
not solely based on the information gathered from the current trial, but also modified by
the information induced by the instructional manipulation and the block context within
which search displays are presented (see also Chun & Jiang, 1999; Sobel & Cave, 2002;
see Chun, 2000, for a review). Thus, the current finding supplements our earlier studies
on the distractor-ratio effect, which highlighted the bottom-up controls of attention in a
search task (Shen et al. 2000, 2003; see also Sobel & Cave, 2002).

The current study differed from Bacon and Egeth (1997) in several important aspects.
First of all, Bacon and Egeth (1997) did not separate the effects of instructional and
contextual manipulations. In their study, participants were informed of which distractor
type would be less frequent on most trials and were asked to restrict their search to a
subset of distractors (e.g., attend to red items or attend to vertical items). In contrast,
in the present study, a congruency effect was observed both with and without explicit
instructions. This finding shows that top-down controls of visual attention come in various
forms (explicit instruction vs implicit knowledge; Wolfe et al., 2003). Similar to previous
demonstrations of implicit top-down guidance by stimulus features (Wolfe et al., 2003),
location context cueing (Chun & Jiang, 1999), knowledge on the probability distribution of certain display types in a group of trials can also induce a bias in deployment of attention during the search process. Future studies may further examine how the guidance of attention and saccadic bias can be modified on a trial-by-trial basis (see Maljkovic & Nakayama, 1994) instead of based on block statistics reported here. The current study also revealed a stronger congruency effect in the instruction condition than in the uninstruction condition. This interaction clearly demonstrates that the presence of top-down influences associated with instructional manipulation operates over and above any influences of implicit block context.

Second, the current study also revealed the spatiotemporal dynamics of visual search behavior by examining the patterns of eye movements accompanying the search process. We found that the congruent trials yielded fewer fixations and shorter fixation durations than did the incongruent trials, whereas the initial saccadic latency did not differ between these two types of trials. In addition, saccadic bias towards the same-color distractors in the color-search displays and towards the same-shape distractors in the shape-search displays was stronger in the congruent trials than in the incongruent trials.

Findings from these oculomotor measures have implications for the study of visual attention. For example, the guided-search model (Cave & Wolfe, 1990; Wolfe, 1994) argues that in a search task, a preattentive parallel process guides the subsequent serial shift of attention through display items. In our earlier studies on the distractor-ratio effect (Shen et al., 2001, 2003), we reported a quadratic change in initial saccadic latency as a function of distractor ratio, with a longer initial saccadic latency in those displays with an approximately 1:1 distractor ratio. It may be speculated that the stronger activation peak associated with extreme distractor ratios results in faster initial saccades. Alternatively, it is possible that the time required for extracting an activation map vary with distractor ratios. In either case, this suggests that the initial saccadic latency is strongly influenced by the bottom-up characteristics of search displays. The current finding of no difference in the initial saccadic latency between the congruent and incongruent trials indicates that participant’s initial saccade targeting was not influenced by the informativeness of the top-down information. However, this does not necessarily mean that providing extra instruction or contextual manipulations does not help the initial segmentation of search displays. On the contrary, other eye movement measures indicate that top-down controls modify the search behavior in a later stage of search. The availability of accurate top-down information makes the rejection of distractor items easier, yielding a shorter fixation duration. It also leads to a higher bias in the selection of the relevant subset of distractors. As a result, fewer fixations are necessitated to make a search decision than in those trials in which misleading information is provided. Such a dynamic picture is hard to capture by a study that focuses on manual response only. The temporal dynamics revealed in the current study are useful for the development of the guided search theory (Cave & Wolfe, 1990; Wolfe, 1994) with respect to the interaction between the time course of the preattentive stage and the serial stage of processing.

Finally, examination of eye movement patterns reveals that the top-down controls of visual attention (instruction and contextual manipulation) do not completely override the
bottom-up contributions. In both the congruent and incongruent trials, saccades were still biased towards the smaller subset of distractors; such a bias was stronger in those trials that were consistent with the block context or instructional manipulations. Thus on inconsistent trials, even with explicit instructions, participants did not stick to the search strategies slavishly by searching entirely through a much larger subset of distractors. Instead, the top-down controls of visual attention were applied flexibly, weighting towards the stimulus dimension that is beneficial to an efficient search. This suggests that human visual search behavior is adaptive and modified flexibly to accommodate the changes to the environment and current task demand (see also Pomplun et al., 2001).

Acknowledgements

Preparation of this manuscript was supported by a grant to Eyal M. Reingold from the Natural Science and Engineering Research Council of Canada (NSERC). We wish to thank Martin Fischer and Adam Galpin for their helpful comments on an earlier draft of the current manuscript.

References


Chapter No: 28

Query No Query

AU1 ‘Wolfe, Butcher, Lee, & Hyle, 2003’ has not been listed in the references. Please check and provide.

AU2 “Therefore, it remains... are changed owing to the... top-down controls” has been changed to “Therefore, it remains... are changed owing to the... top-down controls”. Is this OK?

AU3 ‘Chun & Jiang, 1998’ has been changed to ‘Chun & Jiang, 1999’ in order to match with the references. Please check and provide. Also shall we change in all other occurrences.

AU4 “The temporal dynamics...the interaction between and time course of the... processing” has been changed to the “The temporal dynamics ... the interaction between the time course of the... processing”. Is this OK?

AU5 Please Update.

AU6 Please check if the shortened version of chapter title for running head is OK.

AU7 ‘Wolfe, Butcher, Lee, & Hyle, 1993’ has not been cited in the references. Please check and provide.