Collinear integration affects visual search at V1

Hiu Mei Chow
Department of Psychology, University of Hong Kong, Hong Kong

Li Jingling
Graduate Institute of Neural & Cognitive Sciences, China Medical University, Taichung, Taiwan

Chia-huei Tseng
Department of Psychology, University of Hong Kong, Hong Kong
Department of Psychology, National Taiwan University, Taipei, Taiwan

Perceptual grouping plays an indispensable role in figure-ground segregation and attention distribution. For example, a column pops out if it contains element bars orthogonal to uniformly oriented element bars. Jingling and Tseng (2013) have reported that contextual grouping in a column matters to visual search behavior: When a column is grouped into a collinear (snakelike) structure, a target positioned on it became harder to detect than on other noncollinear (ladderlike) columns. How and where perceptual grouping interferes with selective attention is still largely unknown. This article contributes to this little-studied area by asking whether collinear contour integration interacts with visual search before or after binocular fusion. We first identified that the previously mentioned search impairment occurs with a distractor of five or nine elements but not one element in a 9 × 9 search display. To pinpoint the site of this effect, we presented the search display with a short collinear bar (one element) to one eye and the extending collinear bars to the other eye, such that when properly fused, the combined binocular collinear length (nine elements) exceeded the critical length. No collinear search impairment was observed, implying that collinear information before binocular fusion shaped participants’ search behavior, although contour extension from the other eye after binocular fusion enhanced the effect of collinearity on attention. Our results suggest that attention interacts with perceptual grouping as early as V1.

Selective attention keeps our sensation system from being overwhelmed by the vast amount of information flowing in every minute. We can control our attention focus in a top-down manner, such as looking for someone in the crowd or focusing on exam materials. We are also involuntarily attracted by an irresistible stimulus (i.e., stimulus driven) even when we do not plan to be, such as when observers’ attention is captured by a task-irrelevant distractor in a visual-search task (Turatto & Galfano, 2000, 2001; Turatto, Galfano, Gardini, & Mascetti, 2004). Both top-down and stimulus-driven attentional processes—despite the apparent difference in their nature—reduce our brain’s workload when we efficiently try to make sense of and react to sensory information.

Perceptual grouping is another mechanism that helps us organize the complex visual scene into reduced clusters for processing, according to properties such as proximity, similarity, good continuation, common fate, and closure (Wertheimer, 1924/1938a, 1923/1938b). For instance, collinear integration is a grouping of smaller elements that are oriented in a collinear way (from head to tail) as one object; neurons in the primary visual cortex respond proportionally with the saliency of a collinear contour (W. Li, Piëch, & Gilbert, 2006). The elements in a collinear structure are collectively seen as one contour, and they become easier to detect when they are arranged in a field of random orientations (Field & Hayes, 2004; Field, Hayes, & Hess, 1993; Freeman, Sagi, & Driver, 2001; Hess, Hayes, & Field, 2003).

The relationship between selective attention and collinear integration has been intensively investigated in recent years (Conci, Müller, & Elliott, 2007; Freeman et al., 2001, 2004; Freeman & Driver, 2005; Kimchi, Yeshurun, & Cohen-Savransky, 2007; Ye-
shurun, Kimchi, Sha’shoua, & Carmel, 2008). Freeman, Sag, and Driver (2001) presented observers with a near-threshold central Gabor patch (target) embedded with four Gabor patches at four corners (flankers). Among the four flankers, two were oriented collinear with the central target, while the other two were orthogonal to the central target. It was found that the collinear flankers but not the orthogonal flankers enhanced central target detection—but this collinear facilitation was evident only when participants’ attention was directed to the flankers (to complete a dual Vernier offset judgment). Simply increasing local contrast of the unattended collinear flankers cannot compensate for inattention to generate a comparable effect of collinear integration of attended flankers; this suggests that attention exerts a direct effect on the collinear grouping process instead of an indirect one on local processing (Freeman, Driver, Sag, & Zhaoping, 2003).

Similarly, collinear grouping has also been reported to modulate selective attention. Jingling and Tseng (2013) showed participants a search display of 21 rows × 27 columns containing identical vertical (or horizontal) bars except for a randomly selected column consisting of orthogonal bars. This distractor column was grouped into a collinear (snakelike) or noncollinear (ladderlike) organization (Figure 1). Both types of distractor were salient due to high orientation contrast compared to the background. Participants judged the orientation of a target located on either the distractor column or one of the other columns. Contrary to the hypothesis that salient items capture attention and create search facilitation, Jingling and Tseng found that target search was slower when the target overlapped than when the target did not overlap with the collinear distractor, while there was no difference for a noncollinear distractor. The collinearity-selective impairment persisted when the whole display was rotated by 90° (Figure 1C, D), suggesting that this effect was not due to the global orientation (horizontal vs. vertical) or the background organization (coherent and collinear). Rather, contour grouping (i.e., collinear vs. noncollinear) of the distractor was critical. The search impairment also depended on the collinearity strength of the distractor, as no search interference was observed from a short distractor consisting of only three collinear bar elements. Together, Jingling and Tseng’s results strongly suggest an interaction between collinear integration and attention.

Previous studies have identified the neural networks related to attention (Beck & Kastner, 2005; Corbetta & Shulman, 2002; Kastner & Ungerleider, 2000; Melloni, van Leeuwen, Alink, & Müller, 2012; Serences & Yantis, 2007; Zhang, Zhaoping, Zhou, & Fang, 2012) and collinear integration (Huang, Hess, & Dakin, 2006; Ito & Gilbert, 1999; W. Li et al., 2006); however, the exact site where collinearity interacts with attention is still unknown. According to Z. Li (1999, 2002) and Zhaoping (2005), V1 is a site where neurons respond to collinear integration and other basic features that involve the saliency map for guidance of visual attention. Psychophysical results have shown that contour integration can occur dichoptically with reduced sensitivity, suggesting possible multistage processing of contour integration beyond binocular fusion (Huang et al., 2006). Recent event related potentials (ERP) studies have provided evidence supporting attentional modulation of collinear facilitation at the primary visual cortex (Khoe, Freeman, Woldorff, & Mangun, 2006; Wu, Chen, & Han, 2005), but whether this interaction was behaviorally relevant is not known. The suprathreshold collinear structure in Jingling and Tseng (2013) is a salient area based on the V1 saliency hypothesis, as it contains both high orientation contrast and contour integration. Since an area of high salience leads to high neural activation and performance enhancement, the opposite search impairment associated with this collinear structure is puzzling. It is possible that this behavioral relevance is a result from feedback signals back to V1 from further processing at a higher cortical area. In this study, we asked where the collinearity interacts with selective attention by presenting the search display to observers monocularly or dichoptically. The visual-search performance in these two conditions would suggest whether this interaction is a binocular process.

In Experiment 1, we found that the collinear distractor generates search impairment in a reduced 9 × 9 search display. In Experiment 2, we replicated the distractor-length effect and identified search impairment as present only when the distractor had five or more elements, not when the distractor was a singleton. We used this distinction across different distractor lengths as a basis for later experiments. In Experiment 3, we presented participants through dichoptic viewing, a collinear distractor long enough to impair visual search. If collinear grouping affects visual search after binocular fusion, we expected to see consistent search impairment across all conditions. Monocularly, observers saw a distractor of one, five, and nine elements, which would result in different search behavior if monocular collinear information were used for attentional deployment. In Experiment 4, a monocular control was designed as a baseline comparison, allowing us to quantitatively describe monocular distractor-length dictates for the search effect of collinearity grouping, suggesting that the effect of collinearity on visual search occurs before binocular
Figure 1. (A–D) Visual search display in Jingling and Tseng (2013). (E) A modified search display from Jingling and Tseng (2013) used in Experiment 5B of the current study.
fusion. In Experiment 5, four further control conditions were reported to control for binocular rivalry, imperfect alignment, and imperfect binocular fusion.

**General method**

**Participants**

Fifteen undergraduate students from the University of Hong Kong (HKU) were recruited in Experiment 1. Another 24 HKU students participated in Experiments 2 through 4. Seven, 14, and 24 additional HKU students participated in Experiments 5A, C, and D respectively, and 16 students from China Medical University participated in Experiment 5B. Three extra participants from Experiments 2 through 4 were excluded due to program error. All participants had normal or corrected-to-normal eyesight and were not informed about the purpose of the experiment. They signed consent forms and received course credits for participation.

**Stimuli and apparatus**

For all experiments except Experiment 5B, the stimuli were shown on a 21-in. CRT monitor (ViewSonic), programmed by Matlab with Psychtoolbox Version 3.0.8 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), against a gray background. The screen was divided into two lateral parts, each containing a black-and-white circular frame of 14.0° in diameter with a 9.06° × 9.06° gray square inside to promote fusion between the two eyes when images were projected through a mirror stereoscope. Our mirror stereoscope consisted of a black box of 60 cm (depth) × 50 cm (height) × 55 cm (width), placed in front of the display monitor with a cardboard separator in the middle. Four adjustable mirrors and a chin rest were at the other end of the box, where our participants sat. The total viewing distance was 71.5 cm, which included the distance between the two mirrors for reflection to each eye. A keyboard was used for response collection.

Our search display contained a 9 × 9 regular grid with a unit grid spacing of 1.06° presented inside a circular frame. Homogeneous horizontal (or vertical) white bars (0.16° × 1.04°) were placed in the center of the unit grids (Figures 2 through 5), except for a column of bars (the distractor) that was orthogonal to the rest (background). The location of the distractor column could have a distance of −3, −1, 0, +1, or +3 columns from the center, randomly determined in each trial. One target element bar was broken (0.16°) by a 45° tilt clockwise or counterclockwise from vertical, and its location was at the middle row of −3, −1, 0, +1, or +3 columns. The target column and distractor column were independently determined so that they overlapped 20% of the time.

For Experiment 5B, the search display was composite of 21 × 27 white bars on a black background (Figure 1E). These bars were uniform in orientation except that bars on one column (the distractor, chosen from one of the seven possible columns: the 8th, 10th, 12th, 14th, 16th, 18th, and 20th columns) were oriented at a 90° contrast with the rest of the bars. The distractor column was either 3 bars or 21 bars, and the orientation of the distractor bars was either vertical (collinear, Figure 1A) or horizontal (noncollinear, Figure 1B). Each bar was 18–24 pixels (approximately 0.61° to 0.81° in visual angle) long and 4 pixels wide. The center of the bar was randomly jittered around in a range of +3 to −3 pixels from the center. The target was the same bar broken by a 45° tilt clockwise or counterclockwise at the center of a bar, located at the center (11th) row of 8th, 10th, 12th, 14th, 16th, 18th, and 20th columns. The probability of the target’s being on the distractor column was 1/7; thus the location of the distractor did not predict the location of the target.

**Procedure**

Participants were instructed to view the stimulus display with both eyes open. Before each experiment (except Experiment 5B), they adjusted the stereoscope to align a red rectangle (9.06° × 1.04°) presented to one eye with a black rectangular box of the same size projected to the other eye (Figure 2A).

In all experiments, each trial began with a fixation display (Figure 2B) for 100 ms, followed by a search display. Participants responded as to whether the target gap embedded in the search display was tilted clockwise (right arrow) or counterclockwise (left arrow) by pressing—as quickly as possible while maintaining accuracy—two marked arrow keys on the keyboard. Each trial was terminated when the participant made a response or when 5 s elapsed without response. Accuracy and reaction time (RT) were recorded for analysis.

In Experiment 1, each participant completed 800 trials, for a total of 30 min. In Experiments 2 through 4, another group of participants completed the three experiments of 400–600 trials each in three separate blocks, which lasted for a total of 60 min. In Experiments 5A through D, each participant completed 392–600 trials in each experiment. In all experiments, a break was given every 100 or 150 trials.
Experiment 1: Collinearly organized distractor impaired visual search

This experiment had two modifications from the original visual-search task in Jingling and Tseng (2013) to meet the needs of this current study. First, the search-display grid was reduced from $21 \times 27$ elements to $9 \times 9$ elements. Second, we examined whether monocular or binocular presentation of background elements for the search affected the results. This would be an important basis for later experiments, as it was desirable to have binocular stimuli to enhance fusion.

Participants reported the orientation of the target gap in all four conditions in this experiment. The search display was presented to one eye only (monocular presentation, Figure 3A, D) or so that the other eye also saw the background elements (binocular presentation, Figure 3B, E). The distractor column was organized in either snakelike collinear (Figure 3A, B) or ladderlike noncollinear organization (Figure 3D, E).

Results and discussion

The mean RT from all 15 participants was 717 ms ($SD = 40$ ms). We used only correct trials for RT analysis and excluded the 3.5% of the trials whose RT was above two standard deviations from the grand mean. The RTs for each condition are summarized in Figure 3C and F. The accuracy and RT were submitted to a 2 \(\times\) 2 \(\times\) 2 (Stimulus Presentation [monocular, binocular] \(\times\) Collinearity [collinear, noncollinear] \(\times\) Target Location [overlapping, nonoverlapping]) repeated-measures analysis of variance (ANOVA). The accuracy rates of different conditions are summarized in Table 1.

For RT, there was a significant main effect of stimulus presentation, collinearity, and target location. In general, it took longer to respond in binocular presentation ($M = 715$ ms) than in monocular presentation ($M = 650$ ms), $F(1, 14) = 42.738, MSE = 0.128, p < 0.001$, which likely resulted from the time required for binocular fusion. It took longer to respond to collinear distractors ($M = 714$ ms) than to noncollinear distractors ($M = 652$ ms), $F(1, 14) = 39.741, MSE = 0.117, p < 0.001$, and longer to respond to overlapping targets ($M = 711$ ms) than to nonoverlapping targets ($M = 654$ ms), $F(1, 14) = 31.880, MSE = 0.099, p < 0.001$.

The RT interaction between collinearity and target location is significant: $F(1, 14) = 77.646, MSE = 0.197, p < 0.001$. Post hoc analysis showed that when the distractor was collinear, it took 138 ms longer on average to discriminate overlapping targets ($M = 783$ ms) than nonoverlapping targets ($M = 645$ ms), $t(14) = -7.778$, Holm-adjusted $p < 0.001$. When the distractor was noncollinear, overlapping targets ($M = 640$ ms) enjoyed a small but significant facilitation effect over nonoverlapping targets ($M = 663$ ms), $t(14) = 3.043$, Holm-adjusted $p = 0.025$. We successfully replicated the search-impairment effect with a reduced search display (Jingling & Tseng, 2013).

Our second purpose in this experiment was to examine the effect of monocular and binocular
Figure 3. Stimuli and results of Experiment 1. The distractor refers to the column of bar elements that were orthogonal to the majority of the rest, which could be collinear (snakelike; A, B) or noncollinear (ladderlike; D, E). The distractor and target were presented to one eye randomly, while the background elements were presented either monocularly (on the same side of the collinear distractor and target, as in A and D) or binocularly (on both sides, as in B and E). Targets could overlap with the distractor column (A, E) or not (B, D). Mean response times (RTs) of correct trials for each condition are plotted with the standard error of the mean, indicated by error bars, respectively (C, F). RT for overlapping targets was significantly longer than for nonoverlapping targets on collinear distractors. The search impairment of the collinear distractor and the null search effect of the noncollinear distractor on visual search were consistent in both monocular and binocular presentation of background elements. ** *p < 0.01, *** p < 0.001 (N = 15).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Collinear</th>
<th>Noncollinear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monocular</td>
<td>N/A 97.1 97.4</td>
<td>N/A 97.9 97.9</td>
</tr>
<tr>
<td>Binocular</td>
<td>N/A 97.7 96.4</td>
<td>N/A 96.7 98.6</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>97.8 97.2 98.2 98.5</td>
<td>97.9 96.2 98.2 98.3 N/A N/A</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>97.1 96.2 97.4 97.5 98.5</td>
<td>96.5 96.9 97.8 97.7 97.4 96.8</td>
</tr>
<tr>
<td>Experiment 4</td>
<td>97.5 98 97.6 99.1 N/A N/A 97.9 96.1 98.1 98.5 N/A N/A</td>
<td></td>
</tr>
<tr>
<td>Experiment 5C</td>
<td>96.8 96.2 96.6 96.9 97.6 97.2</td>
<td>97.7 96.9 97.6 96.9 97.4 96.9</td>
</tr>
</tbody>
</table>

Table 1. Mean accuracy rates (%) in experiments where we manipulated monocular-distractor length. Notes: N-Ov: Nonoverlapping target. Ov: Overlapping target.
presentations. In general, regardless of target location, participants spent longer in the binocular presentation of the search display, \( F(1, 14) = 11.635, MSE = 0.024, p = 0.004 \). This suggests that the search-impairment (for a collinear distractor) and facilitation (for a noncollinear distractor) effects are consistent in both monocular and binocular presentations.

ANOVA of the accuracy rate found neither significant main nor interaction effects [collinearity: \( F(1, 14) = 1.429, MSE = 0.001, p = 0.252 \); stimulus presentation: \( F(1, 14) = 0.227, MSE < 0.001, p = 0.641 \); target location: \( F(1, 14) = 0.123, MSE < 0.001, p = 0.731 \); collinearity \( \times \) stimulus presentation: \( F(1, 14) < 0.001, MSE = 6.217 \times 10^{-9}, p = 0.996 \); collinearity \( \times \) target location: \( F(1, 14) = 1.207, MSE = 0.002, p = 0.290 \); stimulus presentation \( \times \) target location: \( F(1, 14) = 0.013, MSE = 1.165 \times 10^{-5}, p = 0.909 \)]}, except the three-way interaction of the factors, which was marginally significant, \( F(1, 14) = 4.566, MSE = 0.002, p = 0.051 \). The longer RT in one condition cannot be explained by a more careful search strategy, which should have generated higher accuracy. This led us to conclude that the RT differences are not due to a trade-off between speed and accuracy. In summary, we validated our stimuli design to probe the original search-impairment effect discovered in Jingling and Tseng (2013).

### Experiment 2: Collinear-distractor size versus search impairment

Jingling and Tseng (2013) reported that collinear grouping strength is crucial to inducing the search impairment: In a \( 21 \times 27 \) element search display, only when a collinear distractor had nine or more elements was the search impairment for overlapping targets significant; no search impairment or facilitation was observed when a collinear distractor had three elements. In the current experiment, we presented distractors of either one or five elements (distractor length: Figure 4A, B) to establish the distractor-length effect. Odd numbers of elements were chosen in order to keep the center of the distractor in the middle row of the whole search display. We include noncollinear distractors (Figure 4D, E) to contrast the results.

### Results and discussion

The mean RT from all 24 participants was 651 ms \((SD = 31 ms; Figure 4C, 4F)\). Trials in which observers’ RT was longer than two standard deviations from the grand mean (approximately 3.4% of the trials) were removed from subsequent analysis. RT analysis included only correct trials. The accuracy and RT data were submitted to a \( 2 \times 2 \times 2 \) (Collinearity [collinear, noncollinear] \( \times \) Distractor Length [1, 5] \( \times \) Target Location [overlapping, nonoverlapping]) repeated-measures ANOVA.

There was no significant main effect of collinearity, distractor length, or target location on RT \((ps > 0.05)\). However, the three-way Collinearity \( \times \) Distractor Length \( \times \) Target Location interaction on RT was significant, \( F(1, 23) = 5.358, MSE = 0.007, p = 0.028 \). Post hoc analysis found that for collinear distractors of only five elements, the RT for overlapping targets \((M = 641 ms)\) was longer than that for nonoverlapping targets \((M = 604 ms)\), \( t(23) = -4.763, Holm-adjusted p = 0.0023 \). There was no search-impairment or facilitation effect when the collinear distractor had only one element or when the distractor was noncollinear \((ps > 0.05)\).

Regarding accuracy, participants were more accurate in responding to trials with distractors of length 5 \((M = 98.3\%)\) than of length 1 \((M = 97.3\%)\), \( F(1, 23) = 12.262, MSE = 0.005, p = 0.002 \), and there was a marginally significant interaction between distractor length and target location, \( F(1, 23) = 3.486, MSE = 0.002, p = 0.075 \). In summary, we did not find evidence for a trade-off between speed and accuracy.

Combining results from Experiments 1 and 2, we found that (a) only collinearly grouped distractors impaired search, (b) this impairment was present only when the distractor had five or more elements, and (c) only the longest noncollinear distractor (nine elements) produced a small facilitation.

### Experiment 3: Collinear search impairment occurred before binocular fusion

In the third experiment, we investigated where in the visual pathway the collinear search-impairment effect was established; this was accomplished by utilizing the effect of distractor length established in Experiment 2. In one eye, targets and collinear distractors of various lengths (monocular-distractor length = one, five, or nine elements) were presented. In the other eye, distractor elements were presented at complementary locations such that when they were fused binocularly, the combined distractor length (i.e., nine elements) always produced search impairment (Figure 5). Background elements were presented in both eyes to promote fusion.
If collinear grouping interacted with visual search only after proper binocular fusion, observers’ search performance should be the same across all conditions—regardless of the difference in monocular-distractor length—because the combined distractor of all conditions is always the same (i.e., nine elements). Otherwise, search performance should be determined by the monocular-distractor information.

Results and discussion

The mean RT from all 24 participants was 755 ms ($SD = 45$ ms). We followed the same criteria as before and discarded 4.4% of the trials. The RT results were plotted (Figure 5). The accuracy and RT data were submitted to a $2 \times 3 \times 2$ (Collinearity [collinear, noncollinear] × Monocular Distractor Length [1, 5, 9] × Target Location [overlapping, nonoverlapping]) repeated-measures ANOVA.

All main effects and interactions on participants’ RTs were significant [collinearity: $F(1, 23) = 49.277$, $MSE = 0.237$, $p < 0.001$; monocular distractor length: $F(2, 46) = 3.484$, $MSE = 0.004$, $p = 0.039$; target location: $F(1,2,3) = 4.778$, $MSE = 0.034$, $p = 0.039$; Collinearity × Monocular Distractor Length × Target Location interaction: $F(2, 46) = 10.586$, $MSE = 0.016$, $p < 0.001$]. A post hoc paired comparison, adjusted with a method of Holm’s sequential Bonferroni, found that at monocular-distractor lengths 5 and 9, the mean RT for overlapping targets was significantly longer than that for nonoverlapping targets, but it was not when the monocular distractor had only one bar element [monocular-distractor length = 5: overlapping targets ($M = 763$ ms) vs. nonoverlapping targets ($M = 674$ ms), $t(23) = -4.903$, Holm-adjusted $p < 0.01$; monocular-distractor length = 9: overlapping targets ($M = 792$ ms) vs. nonoverlapping targets ($M = 672$ ms), $t(23) = -4.043$, Holm-adjusted $p < 0.01$].
$t(23) = -6.332$, Holm-adjusted $p < 0.001$; monocular-distractor length $= 1$: overlapping targets ($M = 717$ ms) vs. nonoverlapping targets ($M = 685$ ms), $t(23) = -1.982$, Holm-adjusted $p > 0.05$]. This pattern resembled those found in Experiments 1 and 2. A separate ANOVA was performed on noncollinear distractors that found that, in general, the RT for overlapping targets ($M = 641$ ms) is shorter than that for nonoverlapping targets ($M = 678$ ms), $F(1, 23) = 9.492$, $MSE = 0.050$, $p = 0.005$, and there was insignificant interaction between monocular-distractor length and target location ($p > 0.05$), suggesting a mild search-facilitation effect among three monocular-distractor lengths.

These effects on RT were not due to a trade-off between speed and accuracy, because all main effects
and interaction effects on accuracy rate were insignificant \((ps > 0.15)\).

In sum, for collinear distractors the search performance was determined by the monocular perceptual collinear information, suggesting that the collinear interference on visual search occurred before binocular fusion. For noncollinear distractors, we found more robust RT facilitation than did previous studies when search arrays were monocularly displayed (Experiment 2), regardless of the monocular-distractor length presented to the eye with the target.

We noted that the monocular display in Experiment 3 was not identical to the search displays in Experiments 1 and 2: The distractor column of the two ends of the distractor was left empty here, while in previous experiments, the rest of the distractor column was filled with orthogonal elements. Thus, we designed Experiment 4 to include a monocular display with distractor columns with blank ends (i.e., identical to Experiment 3) as a control condition to see if the same search pattern persisted.

### Experiment 4: Collinear search impairment persists in monocular display

In this fourth experiment, everything was identical to Experiment 3 except that all search arrays were monocularly presented and the distractor was gapped with blank spaces (Figure 6). If the collinear search-impairment effect was established solely by the monocular image, a pattern of results similar to those in Experiment 3 was expected.

### Results and discussion

The mean RT from all 24 participants was 673 ms \((SD = 34 \text{ ms})\). Criteria to discard data were the same as in previous experiments, and 3.7% of trials exceeding two standard deviations from the grand mean were removed; 2.3% of the trials were incorrect and also did not enter RT analysis, as summarized in Figure 6. The accuracy and RT data for collinear and noncollinear distractors were submitted to a \(2 \times 2\) (Distractor Length \([1, 5]\) \times Target Location \(\{\text{overlapping, nonoverlapping}\}) repeated-measures ANOVA.

For collinear distractors, participants were faster to respond to nonoverlapping targets \((M = 628 \text{ ms})\) than overlapping targets \((M = 665 \text{ ms})\), \(F(1, 23) = 21.411, MSE = 0.024, p < 0.001\). Also, there was a marginally significant interaction effect between distractor length and target location, \(F(1, 23) = 3.250, MSE = 0.003, p = 0.085\). A planned pairwise comparison revealed that when the collinear distractor had only one bar element, the RT for overlapping targets \((M = 644 \text{ ms})\) was not significantly different from that for nonoverlapping targets \((M = 625 \text{ ms})\), \(t(23) = -1.813, \text{Holm-adjusted } p = 0.249\). However, when the collinear distractor had five elements, the RT for overlapping targets \((M = 664 \text{ ms})\) was longer than that for nonoverlapping targets \((M = 620 \text{ ms})\), \(t(23) = -5.331, \text{Holm-adjusted } p < 0.001\). The pattern resembled that of Experiment 3, suggesting that the image presenting a collinear distractor to only one eye is sufficient to provoke search impairment.

For noncollinear distractors, participants were slower to respond to nonoverlapping targets \((M = 636 \text{ ms})\) than to overlapping targets \((M = 620 \text{ ms})\), \(F(1, 23) = 4.764, MSE = 0.006, p < 0.040\). Further, there was a significant interaction effect between distractor length and target location, \(F(1, 23) = 4.785, MSE = 0.003, p = 0.039\). Post hoc analysis revealed that when the noncollinear distractor had only one bar element, the RT for nonoverlapping targets \((M = 644 \text{ ms})\) was significantly shorter than that for overlapping targets \((M = 617 \text{ ms})\), \(t(23) = -3.110, \text{Holm-adjusted } p = 0.025\)—but not when the collinear distractor had five elements \((p > 0.05)\). The distinct pattern of collinear and noncollinear distractor again suggests that search impairment is unique to collinear-distractor lengths of at least five elements.

Regarding accuracy rate, performance was marginally better for length 5 \((M = 98.3\%)\) than length 1 \((M = 97.4\%\), \(F(1, 23) = 3.861, MSE = 0.004, p = 0.062\). There was significant interaction between collinearity and target location, \(F(1, 23) = 4.691, MSE = 0.004, p = 0.041\), and between distractor length and target location, \(F(1, 23) = 4.432, MSE = 0.003, p = 0.046\). Post hoc analysis revealed that individual comparison among the conditions were not significant (Holm-adjusted \(ps > 0.138\)).

### Experiment 5A: Collinear impairment cannot be explained by binocular rivalry

One may note that in all dichoptic displays, the target gap was available to only one eye. Particularly, in the nonoverlapping condition, participants saw the target bar with the oblique gap in one eye and a complete bar without the gap in the other (Figure 3B, 3D, 5B, 5C, 5E, 5G), which could induce possible binocular rivalry. Here we controlled for this possible account by presenting the target gap to both eyes (Figure 7B). This will be compared with the
conditions in Experiment 3 where the nonoverlapping target gap was only shown in one eye (Figure 7A).

**Results and discussion**

The mean RT from seven participants was 702 ms ($SD = 429$ ms). Criteria to discard data were the same as in previous experiments, and 6.2% of trials exceeding two standard deviations from the grand mean were removed. RT analysis is summarized in Figure 7. The accuracy and RT data were entered to a $2 \times 2$ (Target Eye [one eye, both eyes] × Target Location [overlapping, nonoverlapping]) repeated-measures ANOVA.

The same search impairment persisted regardless of target-eye condition [overlapping target: $M = 771$ ms; nonoverlapping target: $M = 626$ ms; $F(1, 6) = 23.138$, $MSE = 0.114$, $p = 0.003$]. Observers were faster when targets were present in both eyes [both eyes: 665 ms; one eye: 702 ms; $F(1, 6) = 24.585$, $MSE = 0.009$, $p = 0.003$]. There was also significant interaction between target location and target eye, $F(1, 6) = 23.061$, $MSE = 0.011$, $p = 0.003$. Post hoc analysis revealed that search impairment with the target on both eyes was greater than with the target on one eye [both eyes: 582 ms vs. 749 ms, Holm-adjusted $p = 0.0037$; one eye: 658 ms vs. 746 ms, Holm-adjusted $p = 0.0450$].

Our results refute that binocular rivalry is a confounding factor for search impairment at targets overlapping with collinear columns.

**Experiment 5B: Collinear impairment cannot be explained by misalignment of the distractor**

In our critical conditions in Experiment 3, a few participants reported that the central part of the distractor (presented to one eye) was not always perfectly aligned with the two ends of the distractor (presented to the other eye). Although we had observers readjust the stereoscope mirrors at the
beginning of each experiment and allowed a 100-ms fusion period at the beginning of each trial, inevitable eye-gaze shift could cause a slight misalignment of the images. To study whether perfect alignment of the collinear elements is needed to generate the search impairment, we conducted a control experiment using the display in Figure 1 except that all bars in the display had random one- to three-pixel offsets from their regular positions (Figure 8A).

Results and discussion

The same discard criteria were applied in this experiment. The selected RTs (Figure 8B) were then submitted to a three-way repeated-measures ANOVA with factors of distractor size (3 or 21 bars), distractor type (collinear or noncollinear), and target type (overlapping or nonoverlapping). Results showed that responses were faster for trials with noncollinear distractors than with collinear distractors, $F(1, 15) = 5.76, MSE = 8,016.49, p < 0.05$, and faster for trials with short distractors than with long distractors, $F(1, 15) = 13.98, MSE = 5,918.68, p < 0.01$. More interestingly, significant two-way interactions were found between distractor type and target type, $F(1, 15) = 40.95, MSE = 8,892.61, p < 0.0001$, and between distractor size and target type, $F(1, 15) = 25.40, MSE = 6,145.13, p < 0.0001$. Further analysis of the simple main effect showed that for trials with collinear distractors, responses to overlapping targets (965.99 ms) were slower than...
those to nonoverlapping targets (856.41 ms), $F(1, 30) = 22.86$, $MSE = 9,027.58$, $p < 0.0001$, while for trials with noncollinear distractors, responses to overlapping targets (825.82 ms) were faster than those to nonoverlapping targets (924.66 ms), $F(1, 30) = 17.64$, $MSE = 9027.58$, $p < 0.0001$. When distractor size was 3, responses were faster to overlapping targets (837.41 ms) than to nonoverlapping targets (900.06 ms), $F(1, 30) = 8.28$, $MSE = 7653.84$, $p < 0.01$. However, when the distractor size was 21, responses were slower to overlapping targets (953.70 ms) than to nonoverlapping targets (880.67 ms), $F(1, 30) = 12.31$, $MSE = 7653.84$, $p < 0.001$. Thus, we were able to find the unique search impairment by long collinear distractors even when the search display was jittered and the alignment between bars rendered imperfect. This suggests that our reported results remain valid with consideration of possible imperfect alignment.

Experiment 5C: Collinear impairment cannot be explained by imperfect binocular fusion

An alternative explanation of the null effect of distractor length 1 in Experiment 3 might be that imperfect binocular fusion led to a misaligned, therefore popped-out, singleton, which canceled out the search-impairment effect by a binocularly fused full distractor. In this experiment, we repeated Experiment 3 with a more thorough procedure to ensure proper binocular fusion, with two additional measurements: (1) Participants initiated each trial after they properly aligned a central Nonius fixation and the outer ring (Figure 9A, B); and (2) to control vergence, each eye only saw a partial outer ring (i.e., a complete outer ring was visible only through proper fusion) throughout the trial of the search display. It was expected that if the search impairment were caused by imperfect binocular fusion, a procedure that promoted better binocular fusion would diminish or alter the search-impairment pattern. One author and 17 observers who were not informed of its purpose participated in this experiment.

Results and discussion

The mean RT from all 18 participants was 735 ms ($SD = 376$ ms). Criteria to discard data were the same as in previous experiments, and 3.4% of trials exceeding two standard deviations from the grand mean were removed; 2.8% of the trials were incorrect and did not enter RT analysis. The RT results are summarized in Figure 9C and D. The accuracy and RT data for collinear and noncollinear display were submitted to a $3 \times 2$ (Monocular Distractor Length [1, 5, 9] × Target Location [overlapping, nonoverlapping]) repeated-measures ANOVA.

For collinear distractors, participants responded more slowly to overlapping targets ($M = 759$ ms) than nonoverlapping targets ($M = 662$ ms), $F(1, 17) = 16.179$, $MSE = 0.253$, $p = 0.001$. The effect of monocular-distractor length was also significant, $F(2, 34) = 3.304$, $MSE = 0.011$, $p = 0.049$. However, post hoc pairwise comparisons found no significant difference between the different distractor lengths ($p$s > 0.06). The interaction between monocular-distractor length and target location was significant, $F(2, 34) = 7.906$, $MSE = 0.019$, $p = 0.002$. Post hoc analysis revealed that search impairment increases with monocular-distractor length [monocular-distractor length = 1: overlapping targets ($M = 720$ ms) vs. nonoverlapping targets ($M = 667$ ms), Holm-adjusted $p = 0.129$; monocular-distractor length = 5: overlapping targets ($M = 756$ ms) vs. nonoverlap-
ping targets ($M = 665$ ms), Holm-adjusted $p = 0.042$; monocular-distractor length = 9: overlapping targets ($M = 801$ ms) vs. nonoverlapping targets ($M = 655$ ms), Holm-adjusted $p = 0.005$).

For noncollinear distractors, participants were faster to detect overlapping targets ($M = 636$ ms) than nonoverlapping targets ($M = 669$ ms), $F(1, 17) = 25.303, MSE = 0.028, p < 0.001$. The other unreported effects were all above the level of marginal significance ($ps > 0.1$). Data analysis on accuracy in both collinear and noncollinear conditions found no significant effects ($ps > 0.1$).

With 18 observers, the results of this experiment highly resembled those of Experiment 3. This rules out the possibility that a search-impairment cancellation effect caused by a popping-out singleton in the condition with length 1 due to imperfect binocular fusion can explain the findings in Experiment 3.

**Experiment 5D: Null effect at length 1 is not due to a pop-out singleton**

It is noted that the lack of search impairment in the short length condition could originate from other factors, such as a pop-out singleton or misalignment. It is critical to more vigorously test these alternatives before drawing a final conclusion regarding the possible neural site. Here we jittered the middle part of the distractor by three pixels so that it misaligned with the...
rest of the distractor column and perceptually popped out (jittered-distractor length = one or five elements; Figure 10A, B, D, E). If pop-out and/or misalignment can fully account for the lack of an impairment effect in the singleton case, we would expect to see the same null effect in the jittered length 1 in this experiment (even though the number of distractor elements exceeds the critical length).

Results and discussion

The mean RT for all 24 participants was 686 ms (SD = 378 ms). Criteria to discard data were the same as in previous experiments, and 3.9% of trials exceeding two standard deviations from the grand mean were removed; 2.1% of the trials were incorrect and did not enter RT analysis, as summarized in Figure 10C and F. The accuracy and RT data for collinear and noncollinear distractors were entered to a 2 × 2 (Jittered Distractor Length [1, 5] × Target Location [overlapping, nonoverlapping]) repeated-measures ANOVA.

For collinear distractors, participants were faster to respond to nonoverlapping targets (M = 622 ms) than overlapping targets (M = 700 ms), \(F(1, 23) = 70.956, \text{MSE} = 0.144, p < 0.001\). Participants were also faster when the jittered-distractor length was 1 (M = 656 ms) than when it was 5 (M = 667 ms), \(F(1, 23) = 5.439, \text{MSE} = 0.003, p = 0.029\). Also, there was a marginally significant interaction effect between jittered-distractor length and target location, \(F(1, 23) = 4.236, \text{MSE} = 0.003, p = 0.051\). A planned pairwise comparison revealed that for both jittered-distractor lengths, the search impairment for overlapping targets was significant (jittered-distractor length = 1: 623 ms vs. 688 ms, Holm-adjusted \(p < 0.001\); jittered-distractor length = 5: 622 ms vs. 712 ms, Holm-adjusted \(p < 0.001\)). For noncollinear distractors, none of the effects on RT were significant (\(ps > 0.3\)). For both collinear and noncollinear distractors, no significant effects on accuracy rate were found (\(ps > 0.3\)), suggesting that there is no accuracy trade-off.

We found that a misaligned singleton was not enough to compensate for the search impairment by a full-length distractor. This provides additional evidence to support the view that the search impairment indeed originates from a monocular site and that the lack of impairment at distractor length 1 in Experiment 3 was
likely due to weak collinearity rather than other factors (e.g., pop-out or misalignment).

**Search-impairment index (SI) comparison across experiments**

To understand the effect of search impairment better, we compared the search impairment between experiments. Two questions were asked for this cross-experiment comparison: (1) whether monocular information alone can account for the search-impairment size as observed or whether information from the other eye also enhances the effect (comparing Experiments 3 and 4), and (2) whether there is some cancellation effect of the misaligned singleton, although one that does not change the RT pattern (comparing Experiments 3 and 5C). To compare the size of search impairment among experiments, a search-impairment index (SI) was computed. Each observer’s RT difference between nonoverlapping and overlapping targets in condition j was divided by that observer’s mean RT in that condition to generate a search-impairment index $SI(i,j)$. When SI was positive, it indicated a search disadvantage from the target’s overlapping with a distractor column; when SI was negative, it indicated a search facilitation. An average SI ($\overline{SI}$) for experiment condition j [$\overline{SI}(j)$] was obtained by averaging the SI of all participants.

$$ SI(i,j) = \frac{RT_{overlapping}(i,j) - RT_{nonoverlapping}(i,j)}{RT(i,j)} \times 100\% $$

(1)

$$ \overline{SI}(j) = \frac{\sum_{i=1}^{N} SI(i,j)}{N} $$

(2)

The average SI ($\overline{SI}$) in Experiments 1 through 4 and 5C are listed in Table 2.

**Results and discussion**

To quantify the effect sizes from monocular and binocular display, we compared the SI in Experiments 3 and 4. We found that when the distractor was similar to a singleton (i.e., one element), the SI did not differ between Experiment 3 (binocular) and Experiment 4 (monocular) [collinear distractors: $t(23) = 0.623, p = 0.539$; noncollinear distractors, $t(23) = 0.935, p = 0.359$]. However, when collinearity strength increased to five elements, we observed significantly greater impairment (in the collinear-distractor condition) and facilitation (in the noncollinear-distractor condition) when additional collinear information was available from the other eye (Experiment 3) [collinear distractors: $t(23) = 3.206, p = 0.004$; noncollinear distractors: $t(23) = -2.88, p = 0.009$]. Although monocular information (Experiment 4) was enough to reveal consistent trends observed from binocular information (Experiment 3), the SI difference suggests that additional grouping information from the other eye enhanced contour integration and its effects on selective attention.

To understand whether a misaligned singleton can possibly cancel out some of the search-impairment effect in Experiment 3, we compared the SI of Experiment 3 with that of Experiment 5C, where the distractor was better aligned because of better binocular fusion. A two-sample $t$ test revealed that at collinear monocular-distractor length 1, the SI of Experiment 5C ($M = 7.28\%$) was not different from that of the same condition in Experiment 3 ($M = 4.18\%$), $t(40) = 0.923, p = 0.359$. The comparison of all other conditions also revealed no difference ($ps > 0.1$). This suggests that we observed not only a similar RT pattern but also a similar search-impairment size in Experiments 3 and 5C, thus rejecting misalignment as a confounding factor for the null effect of distractor-length 1 in Experiment 3.

**General discussion**

In this study, we investigated whether the effect of collinear grouping on attention occurs before or after binocular fusion. We successfully replicated the previous search impairment of a target overlapping with a global distractor formed by collinear elements (Jingling & Tseng, 2013) with a reduced $9 \times 9$ search display in Experiment 1 and the distractor-length effect in Experiment 2. We found that the effect of collinear integration on attention occurs before binocular fusion. We conclude this from results in Experiment 3, where we presented different lengths of distractor in one eye and other distractor elements at complementary locations in the other eye. Observers’ search impairment when a collinear distractor was present followed the monocular but not the binocular information, rejecting the hypothesis that the search impairment is built solely after binocular fusion. Interestingly, we also found evidence demonstrating that binocular collinear integration enhances the influence from task-irrelevant distractors. Our control experiments confirmed that the signature search impairment is not due to defects from possible binocular rivalry or imperfect binocular alignment.
The findings together suggest that the search impairment from collinearity happens somewhere in the visual system that processes information related to the visual information’s eye of origin and that combines information from both eyes. This makes V1 the most plausible location: Before V1, there is no contralateral communication between the two eyes; thus, there is no information from the other eye. After V1, there is limited information about which eye the stimulus is coming from (Hubel & Livingstone, 1987; Hubel & Wiesel, 1968; Zeki, 1978).

Our findings partially concur with the V1 saliency map proposed by Z. Li (1999, 2002) and Zhaoping (2005) in two ways: First, all our experiments consistently showed that only a good continued collinear distractor prolonged the target search, not a noncollinear structure of similar high local contrast. The modulation on search behavior is closely related to the collinear strength. Z. Li’s model is the only salience model that includes the contextual influence of collinear facilitation in addition to other basic visual features for local-contrast computation. This makes it the most relevant model to which we can relate our results. Second, our results converged to suggest V1 as a potential site where collinearity interacts with selective attention, which is supported by other psychophysical and neuroimaging studies (Beck & Kastner, 2005; Khoe et al., 2006; Koene & Zhaoping, 2007; Melloni et al., 2012; Wu et al., 2005; Zhang et al., 2012; Zhaoping & May, 2007). Particularly, Zhaoping (2008) has found that information from one eye—an ocular singleton—dictates the search effect even when it is embedded into the background presented from the other eye when properly fused.

However, a V1 saliency model considers local orientation contrast and collinear computation, predicting that our collinear distractor defines the most salient area in the display, followed by the noncollinear distractor and nondistractor columns. Contradictory to our intuition, observers’ search was most harmed when the target appeared at the most salient area. One speculation is that the collinear distractor forms an object automatically, which delays disengagement of attention to a local target. Previous research (Kimchi et al., 2007; Yeshurun et al., 2009) has found that when the target is not a constituting element of a “perceptual” object, collinearity and closure can induce a perceptual object that automatically draws attention and enhances detection of the target inside the object. In both of those studies, the search display consisted of randomly organized L-shaped elements in different orientations; when some elements were arranged in a collinear and enclosing way, they formed an object. Observers were asked to find the target, which could be outside or inside the object, and to identify its color (Kimchi et al., 2007) or to do a Vernier judgment (Yeshurun et al., 2009). The researchers found that a target inside a square formed by arranging L-shaped elements is detected faster than when a target outside the distractor or a target with no distractor. This suggests an attentional-facilitation effect of a perceptual object automatically induced by collinear/closure perceptual-integration principles. Our search display differed from the display in those studies, as our target gap was part of the distractor. It is possible that our collinear distractor formed a perceptual object, thus making the overlapping target more difficult to segregate from the column and thereby generating the search impairment. Future research can further explore the role of objecthood on such interaction between perceptual grouping and attention.

Our findings also highlight the need to incorporate perceptual grouping into current salience-based attention models (Itti & Koch, 2000, 2001; Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989). Most models of attention differentiate two stages during a typical visual-search task: Primitive features are independently analyzed by individual feature detectors first, and at the second stage, multiple features of the stimuli are bound in conjunction, which is especially important for a target defined by multiple properties. Our study reported that perceptual-grouping principles such as collinearity play a role in visual search, and it differs from these two stages because the feature of the

<table>
<thead>
<tr>
<th>Experiment 1 (N = 15)</th>
<th>Collinear</th>
<th>Noncollinear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monocular</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Binocular</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Monocular-distractor length (bars)</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Experiment 2 (N = 24)</td>
<td>0.39 (7.65)</td>
<td>6.22 (6.60)</td>
</tr>
<tr>
<td></td>
<td>4.18 (10.63)</td>
<td>16.97 (14.81)</td>
</tr>
<tr>
<td>Experiment 3 (N = 24)</td>
<td>2.48 (7.95)</td>
<td>7.35 (5.74)</td>
</tr>
<tr>
<td></td>
<td>7.28 (10.79)</td>
<td>11.87 (13.79)</td>
</tr>
</tbody>
</table>

Table 2. Average search-impairment index (SI; %) and standard deviations (in parentheses) for conditions in different experiments.
neighboring elements has to be taken into consideration, which usually goes beyond the size of a typical receptive field. Either a separate stage (in addition to the conventional two stages) or a further processing based on single-dimension features (i.e., orientation for collinearity, in this case) is required. In any case, our findings suggest that such perceptual-grouping feature maps need to be placed before binocular fusion.

One may note that when the distractor length is one bar only, it is an “ocular singleton” during dichoptic viewing (Experiment 3). Previous research has shown that an ocular singleton can capture attention even if participants are not aware of it (Zhaoping, 2008). This may explain why it takes longer to search for a target in binocular display (Experiments 1 and 3) than in monocular display (Experiments 2 and 4), because observers may have to be freed from the attention-capturing ocular singleton to search among the rest of the scene. This may in turn account for the greater search-impairment index (SI) in Experiment 3 (binocular) than in Experiment 4 (monocular). Because attention captured by an ocular singleton should be the same across collinear and noncollinear conditions, contextual grouping information is still the most critical factor for our observed phenomenon.

The current study leaves at least three possible future research questions. First, it is assumed in the current study that distractors with different collinear strengths (one, five, and nine elements) operate under the same mechanism to affect visual search. However, the uniquely different patterns at one element—which is a singleton that does not require contour integration—could guide attention by a different mechanism. Interestingly, similar studies employing a task-irrelevant singleton have usually revealed search enhancement by the attention-grabbing distractor (Turatto et al., 2004; Turatto & Galfano, 2000, 2001). We were able to observe only occasional facilitation when the singleton distractor was a horizontal bar (i.e., the noncollinear condition). Future research can further contrast search performance with singleton and integrated distractors to understand how they differ in guiding attention distribution. Second, although we found that monocular collinear information dictated visual search—leading to our conclusion that the collinear contour interacts with attention before binocular fusion—exactly how the information at V1 is used is not known. For example, Khoe et al. (2006) have found that the attentional modulation of lateral interactions was manifested at longer-latency changes from the visual cortex but not short-latency changes. It is possible that the modulation of collinear integration on attention can also be referred back to V1 through a feedback mechanism from the higher cortical area. Third, our observers were aware only of the combined perception, while the visual search performance was more sensitive to the monocular-distractor length—of which the observers were not aware—which suggests a dissociation between attention and awareness (Jiang, Costello, Fang, Huang, & He, 2006; Koch & Tsuchiya, 2007). How subconscious perceptual-grouping information influences attention allocation is still a mystery. Future research can further elucidate whether awareness is required in this kind of effect of perceptual grouping on selective attention.

Keywords: visual search, perceptual grouping, collinearity, binocular fusion, V1

Acknowledgments

This work was supported by NSC101-2410-H-039-001-MY2 to LJ as well as the General Research Fund from the Research Grants Council of Hong Kong, China, and the Hong Kong University Seed Funding Programme for Basic Research to C-HT. LJ was supported by the HKU China Affair Office for a visitorship during the preparation of the manuscript. We thank Matt Oxner for technical support and Leon Leong, May Yeung, and Sunny Lee for their assistance in data collection.

Commercial relationships: none.
Corresponding author: Chia-huei Tseng.
Email: CH_Tseng@alumni.uci.edu.
Address: Department of Psychology, University of Hong Kong, Hong Kong.

References


