BRIEF REPORT

Understanding the collinear masking effect in visual search through eye tracking

Janet H. Hsiao[1,2] · **Antoni B. Chan**[3] · **Jeehye An**[1] · **Su-Ling Yeh**[4] · **Li Jingling**[5]

Abstract

Recent research has reported that, while both orientation contrast and collinearity increase target salience in visual search, a combination of the two counterintuitively masks a local target. Through eye-tracking and eye-movement analysis with hidden Markov models (EMHMM), here we showed that this collinear masking effect was associated with reduced eye-fixation consistency (as measured in entropy) at the central fixation cross prior to the search display presentation. As a decreased precision of saccade landing position is shown to be related to attention shift away from the saccadic target, our result suggested that the collinear masking effect may be related to attention shift to a non-saccadic-goal location in expectation of the search display before saccading to the central fixation cross. This attention shift may consequently interfere with attention capture by the collinear distractor containing the target, resulting in the masking effect. In contrast, although older adults had longer response times, more dispersed eye-movement pattern, and lower eye-movement consistency than young adults during visual search, the two age groups did not differ in the masking effect, suggesting limited contribution from ageing-related cognitive decline. Thus, participants' pre-saccadic attention shift prior to search may be an important factor influencing their search behavior.

Keywords Eye movements · Visual search · Attention capture · Hidden Markov Models

Introduction

In visual search, saliency facilitates attentional priority and enhances search performance (Zhaoping, 2014). Saliency can be created by contrast of features from neighboring items, such as a vertical bar among horizontal ones (orientation contrast). Perceptual grouping may further increase salience. According to the V1 model of salience (Zhaoping, 2014),
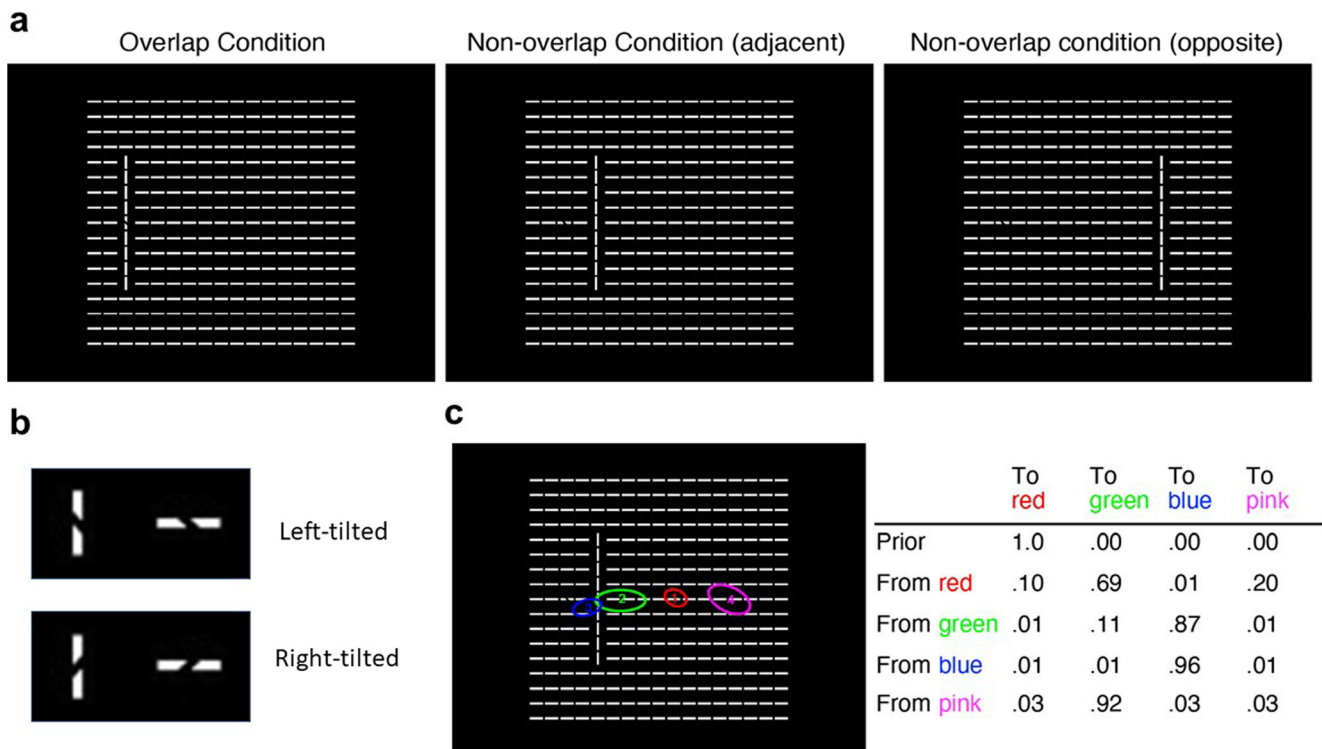
✉ Li Jingling
jlli@mail.cmu.edu.tw

[1] Department of Psychology, University of Hong Kong, Hong Kong, Hong Kong

[2] State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong, Hong Kong

[3] Department of Computer Science, City University of Hong Kong, Hong Kong, Hong Kong

[4] Department of Psychology, National Taiwan University, Taipei, Taiwan

[5] Graduate Institute of Biomedical Sciences, China Medical University, No.91, Hsueh-Shih Road, Taichung 40402, Taiwan

smooth collinear alignment increases salience, making collinear items more salient than non-collinear items (Jingling & Zhaoping, 2008). Intuitively, saliency created by either orientation contrast or collinearity should capture attention and enhance search when the target overlaps with salient items as compared with when it does not (e.g., Turatto & Galfano, 2000). Nevertheless, saliency created by a combination of orientation contrast and collinearity counterintuitively masks a local target, which is referred to as the collinear masking effect (Jingling et al., 2013a, b; Jingling & Tseng, 2013; Tseng & Jingling, 2015; Jingling et al., 2017). This suggests that factors other than saliency may also influence visual search. Here we aimed to investigate potential factors through examining the mechanism underlying the collinear masking effect with eye tracking.

Figure 1a shows a typical search display of the collinear masking effect. Participants search for a bar with a gap and discriminate whether the gap is left- or right-tilted (Fig. 1b). The effect refers to the phenomenon that participants' response times (RTs) are longer when the target is in the collinear column (Fig. 1a, overlap condition) than when it is in the background (Fig. 1a, non-overlap conditions). Although collinearity is shown to increase perceptual salience (Dakin & Baruch, 2009),

Fig. 1 (**a**) Example stimuli in overlap, non-overlap adjacent, and non-overlap opposite conditions. (**b**) Examples of possible targets in the search display. (**c**) An example of an individual hidden Markov model (HMM) trained for one stimulus. Ellipses show regions of interest (ROIs) as two-dimensional Gaussian emissions. The table on the right shows transition probabilities among the ROIs. Priors show the probabilities that a fixation sequence starts from the ellipse

collinearity decreases perceptual salience of orientation contrast, resulting in increased RTs in visual search (Jingling et al., 2013b). Interestingly, other features such as luminance or color do not induce a similar masking effect (Chow et al., 2013; Jingling et al., 2013a). Also, this effect is not due to salience competition (Jingling et al., 2017) or size competition (Chiu & Jingling, 2014) between the target and the collinear distractor, or top-down strategies such as feature-based contingency or practice (Tseng & Jingling, 2015).

Examining individual difference in the collinear masking effect may help us understand why it occurs. Indeed, previous studies have reported a large variance in the effect among individuals, and different cognitive processes may be involved between individuals with and without a masking effect. In addition, since ageing increases motor planning difficulties (Reuter-Lorenz & Park, 2010) and compromises executive attention abilities including distraction suppression (Madden et al., 2014) and top-down attention maintenance (Gazzaley et al., 2005; Hasher & Zacks, 1988), it may enhance the masking effect. Therefore, examining whether older adults have an enhanced masking effect as compared with young adults will help us understand whether the masking effect is related to a decline in executive functions.

Accordingly, here we examined individual differences in the collinear masking effect through eye tracking. In the typical search display (Fig. 1a), the target (a tilted gap on a bar) is small (0.18° by 0.18° of visual angle in the current study) and away from the display center. Thus, eye movement and fixation at the target are required to discriminate the tilt direction (Jingling et al., 2013b). Examining the relationship between one's search behavior in eye movements and strength of the masking effect thus could enhance our understanding of the underlying mechanism. Since the collinear distractor is a super salient item in the display (Jingling & Zhaoping, 2008), participants' eye gaze may be attracted to the collinear distractor initially during search. When the target does not overlap with the collinear distractor (non-overlap condition), participants may continue to search other regions. When the target overlaps with the collinear distractor (overlap condition; Fig. 1a), some participants may detect the target immediately and finish the search, but those experiencing a masking effect may fail to detect the target and continue to search other regions before attending to the collinear distractor again, resulting in more dispersed eye-movement patterns. The masking effect may also increase the uncertainty about when the target can be detected, resulting in less consistent eye-movement patterns across trials. Thus, individual differences in the masking effect are likely to be reflected in eye movements.

To better assess individual differences in eye movements, we used the eye movement analysis with hidden Markov models (EMHMM; Chuk et al., 2014) approach to analyze eye-movement data. EMHMM summarizes a person's eye-

movement pattern in a visual task in terms of both person-specific regions of interest (ROIs) and transitions among the ROIs using a hidden Markov model (HMM), with the optimal number of ROIs determined through a variational Bayesian approach. Thus, it takes individual differences in both spatial and temporal dimensions of eye movements into account. Individual HMMs can be clustered (Coviello et al., 2014) to reveal representative patterns. Differences among models can be quantitatively assessed using data likelihoods, which reflect similarities among individual patterns. Thus, EMHMM provides quantitative measures of individual eye-movement patterns. In addition, we assessed a participant's eye-movement consistency across trials using entropy of the participant's HMM (Cover & Thomas, 2006). Here we examined whether participants who exhibited the masking effect and those who did not differed in eye-movement pattern and consistency during visual search, and whether these effects interacted with participants' age. We hypothesized that participants exhibiting the masking effect may have a more dispersed eye-movement pattern and less consistent eye movements across trials when the target and the collinear distractor were overlapped than when they were not, especially in older adults, due to missing the target, whereas those showing no masking effect may have a more focused eye-movement pattern and consistent eye movements across trials when the target and the collinear distractor were overlapped due to attention capture by salience from collinearity.

Methods

Participants

Forty-four older adults (57–74 years old, M = 66.27; 35 females) and 40 young adults (18–36 years old, M = 22.35; 23 females) were recruited. The sample size was determined according to a similar study comparing young and older adults' eye-movement pattern in a cognitive task using EMHMM (Chan et al., 2018). Here we hypothesized that (1) older adults may have a larger masking effect than young adults (i.e., an interaction between age group and stimulus condition: overlap vs. non-overlap; see Fig. 1a), and (2) individuals with and without a masking effect may have different eye-movement pattern/consistency between the overlap and non-overlap conditions (i.e., an interaction between masking group and stimulus condition; see the Design section for more information). A power analysis showed that the required total sample size for observing a within-between interaction using repeated-measures ANOVA with two groups and two measurements, assuming a medium effect size f = .25, α = 0.05, β = 0.02, is 34. Participants all had normal or corrected-to-normal vision, normal cognitive ability (evaluated by the Mini-Mental State Examination (Folstein et al., 1975), in

which a score over 27 is considered normal), and were able to respond by key pressing. They received honoraria for their participation. Four older adults did not complete the experiment due to eye-tracker problems. The experiment was conducted at the National Taiwan University, and was approved by the Research Ethics Committee of the National Taiwan University.

Materials

A stimulus consisting of horizontal white bars (each extended $0.81° × 0.18°$ of visual angle) arranged in a $17 × 17$ grid (grid spacing $1.04°$ of visual angle) was placed against a black background. The target was a gap ($45°$ tilted, $0.18° × 0.18°$ of visual angle) in the middle of a bar (Fig. 1b). Participants discriminated whether the target was left- or right-tilted in the current search display by key press. The distractor was a set of nine vertical collinear bars. The target and distractor were located in column 3, 5, 13, or 15. The target was at the same height as the fixation cross (row 9), whereas the distractor spanned nine central rows (Fig. 1a; see Design section for more details).

Design

The design consisted of one within-subject variable, stimulus condition (overlap vs. non-overlap), and two between-subject variables, age group (young vs. older adults) and masking group (masking vs. no-masking). The overlap condition consisted of the trials where the distractor and the target were located in the same column (Fig. 1a, left panel). The non-overlap condition consisted of the trials where the distractor and the target were in different columns, including the situation where both the target and the distractor were on either the left or the right side of the display (Fig. 1a, middle panel), or the target and the distractor were on opposite sides of the display (Fig. 1a, right panel). The design for examining the collinear masking effect (Jingling & Tseng, 2013) was under the logic that the distractor was task-irrelevant (e.g., the distractor location did not predict the target location) so that its effect on search performance could be examined independent of task demands (e.g., Turatto & Galfano, 2000; Yantis & Jonides, 1984). Thus, the 16 possible combinations of target and distractor locations were presented with equal frequency. Note that this design resulted in only 25% of the trials belonging to the overlap condition. One may speculate whether the lower probability of the overlap condition could be a confounding factor for the observed colinear masking effect in the literature, i.e., longer RTs in the overlap than the non-overlap condition. This possibility, however, has been ruled out in previous research, where the collinear masking effect could still be observed when the probability of the overlap condition was 60% (Tseng & Jingling, 2015).

We calculated the normalized masking effect as (O - N)/(O + N), where O stands for the response time (RT) in the overlap condition and N stands for the RT in the non-overlap condition; a positive number indicates a masking effect, whereas a negative number indicates otherwise. To understand individual differences, participants were then divided into masking and no-masking groups based on whether they had a masking effect or not. The dependent variables were RT, eye-movement pattern, and eye-movement consistency (as measured in entropy) during the visual search task using the EMHMM method (Chuk et al., 2014), as described in the Eye-movement data analysis section below. A 2 (overlap vs. non-overlap stimulus condition) × 2 (masking vs. no-masking group) × 2 (older vs. young adults) repeated-measures ANOVA was used to examine the effect of stimulus condition, age group, and masking effect group on these dependent variables.

The experiment was conducted using E-prime 2.0 professional with an ASUS BM6660 laptop and a 20-in. i-TECH iF2200 CRT monitor. A standard keyboard was taken as the input device. Eye movement was recorded using Eyelink 1000 plus eye tracker. EyeLink default settings for cognitive research were used for data collection, i.e., saccade motion threshold of 0.1° of visual angle, saccade acceleration threshold of 8,000° per square second, and saccade velocity threshold of 30°. A nine-point calibration procedure was performed in the beginning of the experiment, and recalibration was performed when the average error during validation was larger than 1° of visual angle.

Procedure

Each trial started with a solid dot presented at the center of the screen for drift correction, followed by a fixation point presented at the center of the screen. Participants were told to look at the fixation point when it appeared. After detecting participants' fixation within 1° of visual angle of the fixation point for more than 1,000 ms, a search display was presented at the center of the screen until response. Trials were self-paced and participants pressed a mouse key to start the next trial. Before the experiment, participants were presented with examples of the targets, as shown in Fig. 1b, on an instruction page to ensure they understood that the judgement was based on the orientation of the target/tilted gap. They then used the left and right buttons of the mouse to respond to the left- and right-tilted targets, respectively. An E-prime program was used to record participants' RTs and accuracy. Each participant completed 288 trials of the visual search task, with the 16 stimuli repeating for nine times with the gap on the target bar tilted to the left, and another nine times tilted to the right. The 288 trials were presented in a random order. Participants performed ten practice trials before data collection. The data that

support the findings of this study are openly available at http://dx.doi.org/ 10.17605/OSF.IO/CHKU3.

Eye-movement data analysis

Here we used the EMHMM with co-clustering (Chan et al., 2020b; Hsiao et al., 2021b; Hsiao et al., 2019; see http://visal.cs.cityu.edu.hk/research/emhmm/) to provide quantitative measures of both participants' eye-movement pattern, i.e., where they look (in both spatial and temporal dimensions, or more specifically, eye-fixation locations and transitions among these locations), and eye-movement consistency, i.e., the consistency of where they looked across search trials.

Using EMHMM with co-clustering, each participant's eye movements when viewing one of the 16 stimuli, which was shown for 18 trials during the visual search task, was summarized using an HMM, including person-specific ROIs and eye-gaze transition probabilities among the ROIs. Thus, each participant had 16 HMMs, with each corresponding to an eye-movement pattern of viewing one of the 16 stimuli. As shown in the example in Fig. 1c, the ROIs in an HMM were modeled as two-dimensional (2D) Gaussian distributions. A vector of prior values represented the probability of the first fixation of a trial being located at the corresponding ROI, and a transition matrix indicates the probabilities of fixation transitions among the ROIs. The example shown in Fig. 1c shows that the participant always starts a fixation sequence from the red ROI. Afterwards, the second fixation most likely (69%) lands on the green ROI, and occasionally (20%) lands on the pink ROI. Each HMM was estimated from the individual's data using a variational Bayesian approach to automatically determine the optimal number of ROIs. For each participant's data, HMMs with ROIs ranging from 1 to 6 were trained. To avoid convergence to a local maximum, each HMM was trained 300 times with different initializations. Then, the HMM with the highest log-likelihood of the data was selected.

The machine-learning method co-clustering was used to cluster participants into two groups, group A and group B, according to whether they used similar representative eye-movement patterns across the stimuli. The co-clustering procedure shares the cluster assignments across all the stimuli during the clustering procedure. For each stimulus, the individual HMMs were clustered into two groups according to their similarities using a variational hierarchical EM algorithm (Coviello et al., 2014) to discover common eye-movement patterns among participants. We generated a representative HMM for each common eye-movement pattern. Following previous studies (e.g., Hsiao et al., 2021a; An & Hsiao, 2021; Chuk et al., 2014, 2017a, b; Chan et al., 2018; Zhang et al., 2019), we pre-specified the number of ROIs of the representative HMM as the median number of ROIs among the individual models. Thus, each stimulus had two representative eye-movement pattern HMMs as a result of the

clustering. The co-clustering procedure was conducted 300 times, and the clustering output with the highest data log-likelihood was selected. The resulting two participant groups represented two general eye-movement patterns, pattern A and pattern B, across the stimuli in the visual search task, with each general pattern consisting of a representative HMM for each stimulus. For each participant's eye-movement data, its log-likelihood of being classified as one of the two general eye-movement patterns was calculated as the sum of the data log-likelihoods given the representative HMMs belonging to the general pattern. The log-likelihood indicates the degree of similarity between a person's eye-movement behavior and the general eye-movement pattern discovered through co-clustering. Following previous EMHMM studies (e.g., Chuk et al., 2014; Hsiao et al., 2021b), we tested whether the two general eye-movement patterns were significantly different from each other by examining whether participants using pattern A had a significantly higher data likelihood of pattern A than a data likelihood of pattern B, and whether those using pattern B had a significantly higher data likelihood of pattern B than a data likelihood of pattern A. To quantitatively assess individual eye-movement patterns' similarities along the dimension of the contrast between pattern A and pattern B, following previous studies (e.g., Chan et al., 2018; Chuk et al., 2020; Chan et al., 2020a; Chan et al., 2020c; An & Hsiao, 2021), we defined the A-B scale as below:

$$A-B\ Scale = (A-B)/(|A| + |B|)$$

where A stands for the log-likelihood of the individual eye-movement data being classified as pattern A, and B for the log-likelihood of being classified as pattern B. A more positive A-B scale value indicates a greater resemblance to pattern A. For each participant, the A-B scale for each stimulus was calculated. We then calculated the average A-B scales corresponding to stimuli in the overlap and non-overlap conditions separately for data analysis.

In addition to eye-movement pattern, we assessed each participant's eye-movement consistency across the search trials by calculating the entropy of the participant's eye-movement pattern HMM (Cover & Thomas, 2006), summing over all stimuli. Entropy is a measure of predictability: higher entropy of a participant's HMM indicates a less predictable, less consistent, or more random eye-movement pattern across search trials, and a lower entropy value reflects higher predictability, high consistency, and less randomness. In addition to overall entropy of an HMM, we also calculated the entropy of the first, second, and third fixations (i.e., marginal entropy) during the task separately in order to understand the temporal dynamics of the observed effects. More specifically, overall entropy is calculated from the probability distribution of fixation sequences from the HMM, while marginal entropy is calculated from the probability distribution of first, second,

or third fixations from the HMM. To calculate the marginal entropy of the first fixation of an HMM, we first obtain the probability of viewing each ROI with a first fixation. These probabilities and the corresponding 2D Gaussian ROIs are then used to form the probability distribution of first fixation locations as a Gaussian mixture model (GMM), from which we calculate the entropy. The marginal entropies for the second and third fixations are calculated likewise. Note that here the first fixation of a trial referred to the fixation at the central fixation cross before the search display, which was important for examining effects related to expectation of the search display. Accordingly, the second fixation of a trial was the first fixation landed on the search display after the search display was presented; this fixation was important for examining effects related to attention capture by the salient colinear distractor. The third fixation of a trial was then the second fixation landed on the search display, which was important for examining effects related to target detection or the masking effect. The marginal entropy measure reflected the spread (or variance) of the distribution of each individual fixation. It allowed us to examine the relationship between saccade accuracy and the masking effect. Here we focused our analysis on these first three fixations.

Results

Reaction times

Based on the normalized masking effect, 56 participants (31 young vs. 25 older adults) were classified into the masking group and 24 into the no-masking group (nine young vs. 15 older adults). The numbers of young and older adults classified into the masking and no-masking groups did not differ significantly, $X^2(1, N=80) = 2.25$, $p = .09$. ANOVA (Table A1) on RT with stimulus condition, age group, and masking group as independent variables[1] showed a main effect of age group, $F(1, 76) = 69.36$, $p < .001$, $\eta_p^2 = .48$: young participants responded faster than older participants. There was an interaction between stimulus condition and masking group, $F(1, 76) = 19.40$, $p < .001$, $\eta_p^2 = .20$: the masking group's non-overlap RT was faster than overlap RT, $t(55) = -4.32$, $p < .001$, $d = .58$, indicating a masking effect. In contrast, the no-masking group's non-overlap RT was slower than overlap RT, $t(23) = 2.48$, $p = .021$, $d = .51$, suggesting attention capture (Fig. 2b). This interaction confirmed the masking and no-masking group categorization.

---

[1] In a separate analysis, we conducted two separate 2 × 2 ANOVA to test our two hypotheses, i.e., an interaction between age group and stimulus condition, and an interaction between masking group and stimulus condition, and obtained similar results. See Tables A5 to A10 in the Online Supplementary Materials.
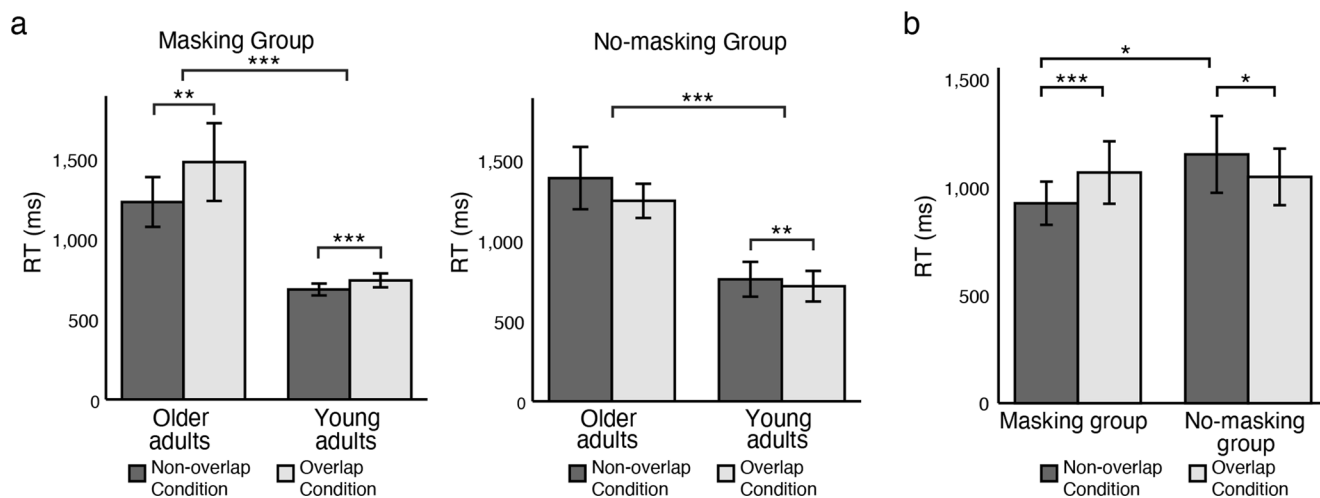
Fig. 2 (**a**) Participants' response times (RTs) in different experimental conditions. (**b**) Interaction effect between stimulus condition and masking group

Eye-movement pattern

EMHMM with co-clustering revealed two representative eye-movement patterns, pattern A and B (Fig. 3, Figs. A1, A2, and A3 (Online Supplemental Material, OSM)). In pattern A, in most stimuli, the eye-movement pattern consisted of smaller and distinct ROIs covering the target and the distractor respectively, or had fixations covering a smaller region of the search display than pattern B. We referred to this pattern as a "concentrated" pattern. In contrast, in pattern B, in most stimuli, the eye-movement pattern consisted of more elongated and overlapping ROIs, or had fixations covering a wider region of the search display than pattern A. We referred to this pattern as a "dispersed" pattern. The two patterns were significantly different, confirming the clustering validity: concentrated patterns had higher data log-likelihood given the concentrated HMMs than dispersed HMMs, $t(48) = 9.40$; $p < .001$, $d = 1.34$, and vice versa for dispersed patterns, $t(30)=14.71$; $p < .001$, $d = 2.27$. Participants using the two patterns did not differ in eye-movement entropy (overall, $p = .70$; marginal, first fixation, $p = .32$; second fixation, $p = .92$; third fixation, $p = .96$). No significant correlation was observed between eye-movement pattern (A-B scale) and entropy (overall, $p = .74$; marginal, first fixation, $p = .61$; second fixation, $p = .77$; third fixation, $p = .60$).

In both patterns, the collinear distractor attracted attention. For instance, when the target and the collinear distractor were on the opposite sides (Fig. 3, middle panels), in the concentrated pattern, participants had a higher probability of shifting gaze from the center to the collinear distractor (red to green, .47) to from the center to the target (red to blue, .35). Similarly, in the dispersed pattern, participants had a higher probability of shifting gaze from the center to the collinear distractor (red to green, .63) than from the center to the target (red to blue, .32). Thus, the collinear distractor attracted
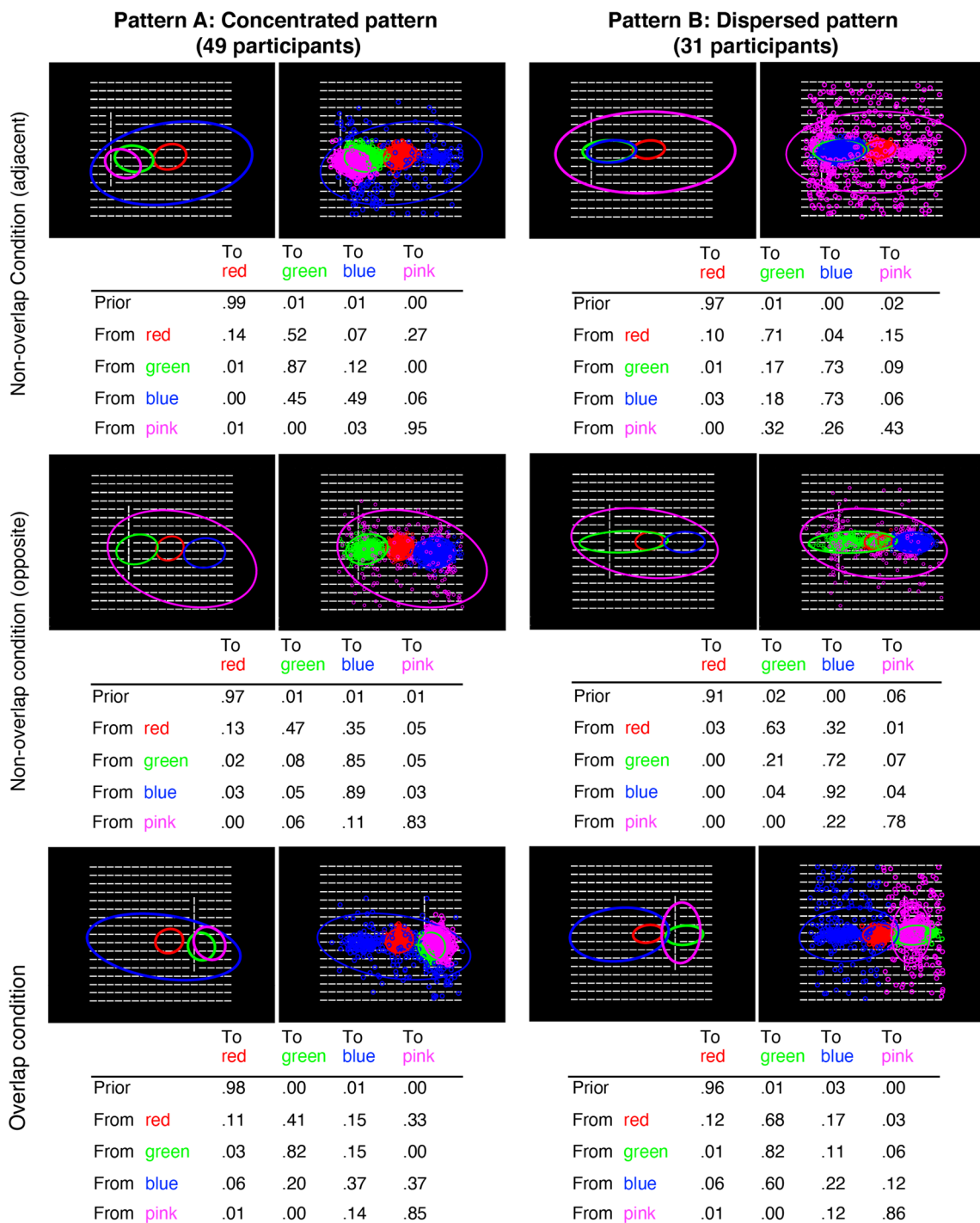
attention either similarly to or more often than the target, suggesting that it was perceptually salient.

ANOVA on A-B scale with stimulus condition, age group, and masking group as variables (data in Table A2, ANOVA details in Table A3 (OSM)) showed a main effect of age group, $F(1, 76) = 9.022$, $p = .004$, $\eta p2 = .11$: young adults ($M = .030$, $SE = .010$) had a higher A-B scale than older adults ($M = -.0086$, $SE = .0084$), suggesting young adults' eye movements were more similar to the concentrated pattern. There was no effect of masking group. Neither the interaction between age group and stimulus condition nor the interaction between masking group and stimulus condition was significant.

Eye-movement entropy

Overall entropy ANOVA (Table A4, OSM) on overall entropy with stimulus condition, age group, and masking group as variables showed a main effect of stimulus condition $F(1, 76) = 28.23$, $p < .001$, $\eta_p^2 = .27$: higher entropy in the non-overlap than overlap condition; and a main effect of age group, $F(1, 76) = 5.29$, $p = .024$, $\eta_p^2 = .065$: older adults had higher entropy than young adults. The main effect of masking group was marginal, $F(1, 76) = 3.80$, $p = .055$, $\eta_p^2 = .048$ (Table 1).

Also, we found an interaction between stimulus condition and masking group, $F(1, 76) = 30.88$, $p < .001$, $\eta_p^2 = .29$ (Table 1): the no-masking group had higher entropy in the non-overlap than in the overlap condition, $t(23) = 6.89$, $p < .001$, $d = .56$, whereas the masking group did not. Since the no-masking group responded faster in the overlap than in the non-overlap condition, suggesting attention capture by the overlapping collinear structure and target, the increased eye-movement consistency in the overlap condition may be related to attention capture. In contrast, in the masking group, this

Pattern A: Concentrated pattern
(49 participants)

Pattern B: Dispersed pattern
(31 participants)

Non-overlap Condition (adjacent)

	To red	To green	To blue	To pink
Prior	.99	.01	.01	.00
From red	.14	.52	.07	.27
From green	.01	.87	.12	.00
From blue	.00	.45	.49	.06
From pink	.01	.00	.03	.95

	To red	To green	To blue	To pink
Prior	.97	.01	.00	.02
From red	.10	.71	.04	.15
From green	.01	.17	.73	.09
From blue	.03	.18	.73	.06
From pink	.00	.32	.26	.43

Non-overlap condition (opposite)

	To red	To green	To blue	To pink
Prior	.97	.01	.01	.01
From red	.13	.47	.35	.05
From green	.02	.08	.85	.05
From blue	.03	.05	.89	.03
From pink	.00	.06	.11	.83

	To red	To green	To blue	To pink
Prior	.91	.02	.00	.06
From red	.03	.63	.32	.01
From green	.00	.21	.72	.07
From blue	.00	.04	.92	.04
From pink	.00	.00	.22	.78

Overlap condition

	To red	To green	To blue	To pink
Prior	.98	.00	.01	.00
From red	.11	.41	.15	.33
From green	.03	.82	.15	.00
From blue	.06	.20	.37	.37
From pink	.01	.00	.14	.85

	To red	To green	To blue	To pink
Prior	.96	.01	.03	.00
From red	.12	.68	.17	.03
From green	.01	.82	.11	.06
From blue	.06	.60	.22	.12
From pink	.01	.00	.12	.86

Fig. 3 Examples of the two general eye-movement patterns (concentrated vs. dispersed) derived by clustering, one each for the adjacent, opposite and overlap stimulus conditions. The small dots show raw fixation locations; the color indicates their region of interest (ROI) assignment. **Upper panel:** The target was adjacent to the collinear distractor. The concentrated pattern showed small and distinctive ROIs for the target (green) and distractor (pink), whereas the dispersed pattern showed relatively larger and overlapping ROIs for the target and distractor (green and blue). **Middle panel:** The target and the collinear distractor were on opposite sides. The concentrated pattern had more distinct, non-overlapping ROIs and a lower probability of being attracted to the distractor relative to the target (.47 vs. .35, from the central fixation cross to the distractor vs. target, respectively) as compared with the dispersed pattern (.63 vs. .32, from the fixation cross to the distractor vs. target, respectively). **Lower panel:** The target and the distractor overlapped. The concentrated pattern had smaller ROIs with the pink ROI more focused on the target area than the dispersed pattern. The raw fixations were more spread out in the dispersed pattern than in the concentrated pattern

Table 1 Descriptive statistics of entropy and marginal entropy measures

	Entropy		M1		M2		M3	
	M	SD	M	SD	M	SD	M	SD
Stimulus condition								
Overlap	7.88	.08	7.50	.08	8.58	.08	8.49	.09
Non-overlap	8.07	.07	7.53	.08	8.92	.07	8.82	.07
Age group								
Young adults	7.80	.11	7.45	.12	8.54	.11	8.37	.12
Older adults	8.14	.10	7.57	.11	8.96	.10	8.94	.10
Masking group								
Masking	8.12	.08	7.68	.09	8.90	.08	8.01	.08
No-masking	7.83	.13	7.35	.14	8.60	.13	8.50	.13

Note: M1 is the marginal entropy of the first fixation; M2 is the marginal entropy of the second fixation, while M3 is the marginal entropy of the third fixation

attention capture did not happen, and eye-movement consistency did not differ between the two conditions (Table 1).

The (normalized) masking effect was positively correlated with overall entropy in the overlap condition, $r(78) = .38$, $p < .001$, but not the non-overlap condition, $r(78) = -.054$, $p = .63$ (Fig. 4a)[2]: the less consistent the eye movements in the overlap condition, the stronger the masking effect.

We also found that overall entropy was positively correlated with RT, $r(78) = .47$, $p < .001$ (Fig. 4b). Since RT was also positively correlated with number of fixations, $r(78) = .39$, $p < .001$, we performed a hierarchical multiple regression analysis to examine whether overall entropy accounted for additional variance in RT after number of fixations was accounted for. At step one, number of fixations was a significant predictor, $R^2 = .53$, $\beta = .73$, $p < .001$. At step 2, adding entropy as a predictor significantly accounted for additional 4.1 % of variance, $\beta = .22$, $p = .009$. Thus, overall entropy accounted for variances in RT independent of number of fixations.

Marginal entropy, first fixation ANOVA with stimulus condition, age group, and masking group as variables (Table A4, OSM) revealed a main effect of masking group, $F(1, 76) = 4.30$, $p = .042$, $\eta_p^2 = .054$: the masking group had less consistent first fixation than the non-masking group (Table 1).

Marginal entropy, second fixation ANOVA with stimulus condition, age group, and masking group as variables (Table A4, OSM) showed a main effect of stimulus condition, $F(1, 76) = 109.92$, $p < .001$, $\eta_p^2 = .59$: participants had higher

entropy in the non-overlap than in the overlap condition; a main effect of masking group, $F(1, 76) = 4.05$, $p = .048$, $\eta_p^2 = .051$: masking group had higher entropy than non-masking group; and a main effect of age group, $F(1, 76) = 8.04$, $p = .006$, $\eta_p^2 = .096$: young adults had lower entropy than older participants. There was an interaction between stimulus condition and masking group, $F(1, 76) = 22.29$ $p < .001$, $\eta_p^2 = .23$ (Table 1): the no-masking group had a more consistent second fixation in the overlap condition than the masking group, consistent with the results in overall entropy.

Marginal entropy, third fixation Similarly, ANOVA with stimulus condition, age group, and masking group as variables (Table A4, OSM) showed a main effect of stimulus condition, $F(1, 76) = 53.32$, $p < .001$, $\eta_p^2 = .41$; a main effect of masking group, $F(1, 76) = 3.97$, $p = .049$, $\eta_p^2 = .050$; a main effect of age group, $F(1, 76) = 13.38$, $p < .001$, $\eta_p^2 = .15$; and an interaction between stimulus condition and masking group, $F(1, 76) = 50.73$ $p < .001$, $\eta_p^2 = .40$.

Discussion

Here we showed that participants' collinear masking effect was reflected in eye-movement consistency across trials, but not eye-movement pattern. Specifically, the masking effect could be predicted by eye-movement consistency in the overlap condition: the less consistent the eye movements, the larger the masking effect (Fig. 4a). In addition, the no-masking group demonstrated more consistent eye movements in the overlap than in the non-overlap condition, suggesting that the attention capture by the overlapping collinear structure and target (as reflected in reduced RT) was associated with enhanced eye-movement consistency. In contrast to eye-movement pattern, which reflected attention capture by where participants looked, eye-movement consistency reflected the spread of where they looked, and thus measured a different aspect of attention capture.

Interestingly, in the temporal dynamics of eye-movement consistency, as compared with the no-masking group, the masking group had significantly higher marginal entropy of the first fixation of a trial, which was aimed at the central fixation cross before the search display. The marginal entropy reflected the variance (or spread) of the first fixation distribution, suggesting that the masking effect was related to the spatial precision of the fixation. This effect did not interact with either age or stimulus condition. In the literature, decreased precision of saccade landing position is shown to be related to attention shift away from the saccadic target (Gersch et al., 2008; Kowler et al., 1995; Wilder et al., 2009; Zhao et al., 2012). Thus, the larger variance of first fixation distribution in the masking group may be due to shift of attention away from the central fixation cross in expectation of the

---

[2] Similar results were obtained when we did robust correlation analysis by winsorizing the masking effect in RT with a 95% confidence interval: A significant positive correlation between masking effect and overall entropy in the overlap condition, $r(78) = .41$, $p < 0.01$.
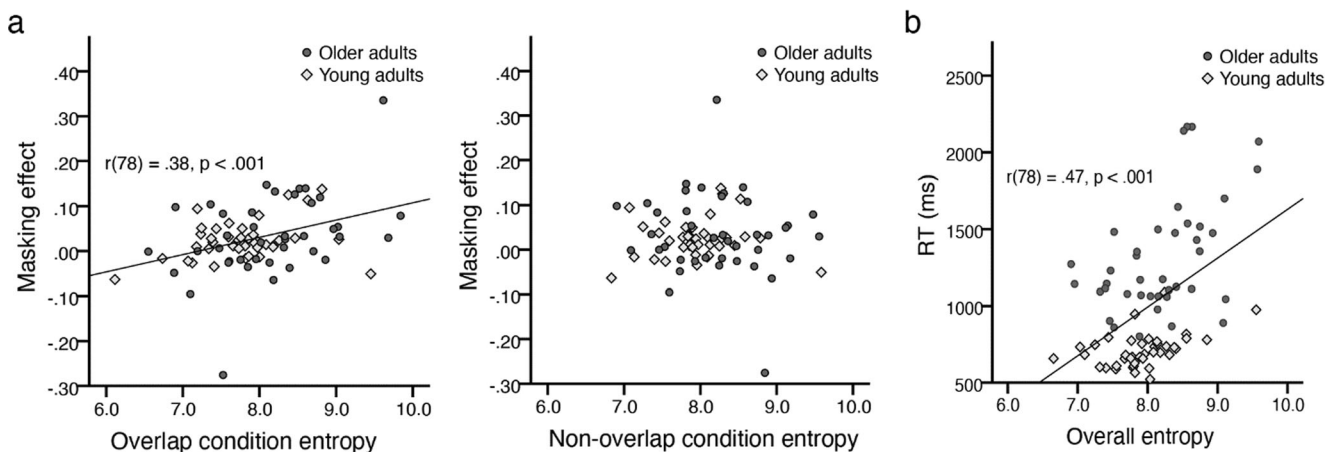
Fig. 4 (**a**) A significant correlation between overall entropy and masking effect in the overlap condition but not in the non-overlap condition. (**b**) Significant positive correlation between overall entropy and response time (RT). This was found in both young, r(38) = .45, p = .004, and older adults, r(38) = .48, p = .002

search display before fixating at the cross. This attention shift may have interfered with saccade accuracy of subsequent eye movements, especially when the collinear distractor did not match the attention shift, resulting in reduced attention capture at the second fixation. Consistent with this speculation, the masking group also had higher marginal entropy of the second fixation than the non-masking group, particularly in the overlap condition, where the collinear distracter with the embedded target was the only salient region in the display. Thus, the masking effect may be caused by a pre-saccadic attention shift to a non-saccadic-goal location that interfered with attention capture by the collinear distractor. This speculation is consistent with Jingling et al.'s (2013b) finding that impulsive saccades associated with perceptual salience were reduced during the first saccade in the overlap condition, suggesting reduced perceptual saliency of the collinear distractor. It is also consistent with previous findings that the masking effect was related to processes taking place as early as V1 (Chow et al., 2013), within 40 ms of search display presentation (Chiu & Jingling, 2015), and not affected by target occurrence probability, target type, practice (Tseng & Jingling, 2015), or salience manipulation (Jingling et al., 2017). Future work may test this possibility through explicit instructions on attention maintenance.

Through clustering, we discovered two significantly different eye-movement patterns (Fig. 3), suggesting that visual search was affected by top-down attention in addition to bottom-up salience (Einhäuser et al., 2008). In both patterns, participants may re-visit the target after viewing the distractor or other areas, suggesting that search can be misled by objects containing the same features (Zhaoping & Guyader, 2007). Interestingly, older adults' eye-movement patterns were more similar to the dispersed pattern than those of young adults, and they had lower eye-movement consistency, consistent with their longer RTs. This phenomenon may be related to

cognitive decline in attention mechanisms. Indeed, ageing can affect both top-down and bottom-up attention during visual search, resulting in longer RTs (Madden, 2007). A similar phenomenon was observed in face recognition (Chan et al., 2018), where older adults' reduced looking at facial features was associated with worse recognition performance and lower visual attention and planning abilities (see also Hsiao et al., 2020). Future work may examine this possibility. Note, however, that despite these observed age differences in RT and eye-movement measures, the two age groups did not differ in the masking effect. This finding suggests that decline in executive functions is unlikely to be a contributing factor to the masking effect.

In our task, there were fewer trials in the overlap than in the non-overlap condition. However, no significant main effect of stimulus condition was observed, consistent with previous research showing that the masked effect is unrelated to the probability of the overlap condition (Tseng & Jingling, 2015). In addition, no significant difference in the masking effect between the two age groups was observed, and the numbers of older and young adults in the masking and no-masking groups did not differ significantly (p = .09). Thus, there was little evidence suggesting that this probability difference could be associated with age difference in the masking effect. Future work may manipulate the probability of the overlap condition and examine whether it induces age difference in the masking effect.

In visual search, changes in predictability of target location may influence eye-movement entropy (Hong & Beck, 2010). Thus, participants' eye-movement entropy may change with task familiarity. To examine this possibility, we divided each participant's search trials into early and late trials, and conducted a 2 (time: early vs. late trials) × 2 (masking group: masking vs. no-masking) ANOVA on overall entropy. We found a main effect of time, F(78) = 45.41, p < .001: eye

movements in late trials were more predictable than early trials. However, this effect did not interact with masking group (p = .62), suggesting that increased predictability of eye movements over time did not account for the difference in masking effect across participants.

In conclusion, here we showed that the collinear masking effect was associated with reduced eye-movement consistency. This phenomenon could be observed in the fixation at the central fixation cross before the search display, suggesting that the masking effect may be related to a pre-saccadic attention shift to a non-saccadic-goal location that is associated with decrease saccade accuracy. This attention shift may consequently interfere with attention capture by the overlapping collinear distractor and target, resulting in the masking effect. In addition, among participants who showed the masking effect, young adults had higher eye-movement consistency when the target and collinear distractor were overlapped than when they were not, suggesting better attention capture, whereas older adults showed the opposite effect. Thus, older adults may have more difficulty recovering from the masking effect than young adults, resulting in an enhanced masking effect. The entropy analysis based on EMHMM makes it possible to reveal these effects, suggesting that participants' pre-saccadic attention shift before search and age-related decline in attention mechanisms may both influence participants' search behavior.

References

An, J., & Hsiao, J. H. (2021). Modulation of mood on eye movement pattern and performance in face recognition. *Emotion, 21(3),* 617-630.

Chan, C. Y. H., Chan, A. B., Lee, T. M. C., & Hsiao, J. H. (2018). Eye movement patterns in face recognition are associated with cognitive decline in older adults. *Psychonomic Bulletin & Review*, *25*(6), 2200-2207.

Chan, F. H. F., Barry, T. J., Chan, A. B., & Hsiao, J. H. (2020a). Understanding visual attention to face emotions in social anxiety using hidden Markov models. *Cognition & Emotion*, *34*(8), 1704-1710.

Chan, F. H. F., Jackson, T., Hsiao, J. H., Chan, A. B., & Barry, T. J. (2020b). The interrelation between interpretation biases, threat expectancies and pain-related attentional processing. *European Journal of Pain, 24*(10)*, 1956-1967.*

Chan, F. H. F., Suen, H., Hsiao, J. H., Chan, A. B., & Barry, T. J. (2020c). Interpretation biases and visual attention in the processing of ambiguous information in chronic pain. *European Journal of Pain*, *24*(7), 1242-1256.

Chiu, C.-W., & Jingling, L. (2014). Focus size of attention: Exploring the focus hypothesis on the collinear masking effect in visual search. *Chinese Journal of Psychology*, *56*(4), 467-484. https://doi.org/10.6129/CJP.20140515

Chiu, C.-W., & Jingling, L. (2015). Temporal characteristics of the collinear masking effect in visual search. *Chinese Journal of Psychology*, *57*(1), 91-107. https://doi.org/10.6129/CJP.20141217

Chow, H. M., Jingling, L., & Tseng, C. H. (2013). Collinear integration affects visual search at V1. *Journal of Vision, 13*(10), 24-24.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *Journal of Vision*, *14*(11):8, 1-14.

Chuk, T., Crookes, K., Hayward, W. G., Chan, A. B., & Hsiao, J. H. (2017a). Hidden Markov model analysis reveals the advantage of analytic eye movement patterns in face recognition across cultures. *Cognition*, *169*, 102-117.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2017b). Is having similar eye movement patterns during face learning and recognition beneficial for recognition performance? Evidence from hidden Markov modeling. *Vision Research*, *141*, 204-216.

Chuk, T., Chan, A. B., Shimojo, S., & Hsiao, J. H. (2020). Eye movement analysis with switching hidden Markov models. *Behavior Research Methods*, *52*(3), 1026-1043.

Cover, T. M., & Thomas, J. A. (2006). Elements of Information Theory (2nd ed.). Interscience.

Coviello E. Chan A. B. Lanckriet G. R. G. (2014). Clustering hidden Markov models with variational HEM. *Journal of Machine Learning Research*, *15* (Feb), 697–747.

Dakin, S. C., & Baruch, N. J. (2009). Context influences contour integration. *Journal of Vision*, *9*(2), 13-13.

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14), 18-18.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198.

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

Gersch, T. M., Kowler, E., Schnitzer, B. S., & Dosher, B. A. (2008). Visual memory during pauses between successive saccades. *Journal of Vision*, *8*(16):15, 1-18.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The psychology of learning and motivation*, *22*, 193-225.

Hong, S. L., & Beck, M. R. (2010) Uncertainty Compensation in Human Attention: Evidence from Response Times and Fixation Durations. *PLOS ONE*, *5*(7), e11461. https://doi.org/10.1371/journal.pone.0011461

Hsiao, J. H., Chan, K. Y., Du, Y., & Chan, A. B. (2019). Understanding Individual Differences in Eye Movement Pattern During Scene Perception through Hidden Markov Modeling. In A.K. Goel, C.M. Seifert, & C. Freksa (Eds.), *Proceeding of the 41th Annual Conference of the Cognitive Science Society* (pp. 3283). Cognitive Science Society.

Hsiao, J. H., An, J. H., & Chan, A. B. (2020). The role of eye movement consistency in learning to recognise faces: Computational and

experimental examinations. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 1072-1078). Cognitive Science Society.

Hsiao, J. H., An, J., Zheng, Y., & Chan, A. B. (2021a). Do portrait artists have enhanced face processing abilities? Evidence from hidden Markov modeling of eye movements. *Cognition*, *211*, 104616.

Hsiao, J. H., Lan, H., Zheng, Y., & Chan, A. B. (2021b). Eye Movement analysis with Hidden Markov Models (EMHMM) with co-clustering. *Behavior Research Methods*, 1-14.

Jingling, L., & Tseng, C. H. (2013). Collinearity Impairs Local Element Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1):156-67. https://doi.org/10.1037/a0027325.

Jingling L., & Zhaoping L. (2008). Change detection is easier at texture border bars when they are parallel to the border: Evidence for V1 Mechanisms of Bottom-up Salience. *Perception*, *37*(2), 197-206.

Jingling, L., Tseng, C.-H., & Zhaoping, L. (2013a). Orientation is different: Interaction between contour integration and feature contrasts in visual search. *Journal of Vision*, *13*(3), 1–13. https://doi.org/10.1167/13.3.26

Jingling, L., Tan, D.-L., & Tseng, C.-H. (2013b). Salient collinear grouping diminishes local salience in visual search: An eye movement study. *Journal of Vision*, *13*(12), 1–10. DOI https://doi.org/10.1167/13.12.6

Jingling, L., Lu, Y. H., Cheng, M., & Tseng, C. H. (2017). Collinear masking effect in visual search is independent of perceptual salience. *Attention, Perception, & Psychophysics*, *79*(5), 1366-1383.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897-1916.

Madden, D. J. (2007). Aging and visual attention. *Current Directions in Psychological Science*, *16*(2), 70-74.

Madden, D. J., Parks, E. L., Davis, S. W., Diaz, M. T., Potter, G. G., Chou, Y. H., … & Cabeza, R. (2014). Age mediation of frontoparietal activation during visual feature search. *NeuroImage*, *102*, 262-274.

Reuter-Lorenz, P. A., & Park, D. C. (2010). Human neuroscience and the aging mind: a new look at old problems. *The Journals of Gerontology: Series B*, *65*(4), 405-415.

Tseng, C. H., & Jingling, L. (2015). A Salient and Task-Irrelevant Collinear Structure Hurts Visual Search. *PLoS ONE*, *10*(4): e0124190.

Turatto, M., & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research*, *40*(13), 1639-1643.

Wilder, J. D., Kowler, E., Schnitzer, B. S., Gersch, T. M., & Dosher, B. A. (2009). Attention during active visual tasks: counting, pointing, or simply looking. *Vision Research*, *49*(9), 1017-1031.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *Journal of Experimental Psychology: Human perception and performance*, *10*(5), 601-621.

Zhang, J., Chan, A. B., Lau, E. Y. Y., & Hsiao, J. H. (2019). Individuals with insomnia misrecognize angry faces as fearful faces while missing the eyes: an eye-tracking study. *Sleep*, *42*(2), zsy220.

Zhao, M., Gersch, T. M., Schnitzer, B. S., Dosher, B. A., & Kowler, E. (2012). Eye movements and attention: The role of pre-saccadic shifts of attention in perception, memory, and the control of saccades. *Vision Research*, *74*(1), 40-60.

Zhaoping, L. (2014). *Understanding vision: theory, models, and data*. Oxford University Press.

Zhaoping, L., & Guyader, N. (2007). Interference with bottom-up feature detection by higher-level object recognition. *Current Biology*, *17*(1), 26-31.