scientific reports

Check for updates

OPEN

Evaluating the contribution of parallel processing of color and shape in a conjunction search task

Andrea Yaoyun Cui, Simona Buetti, Zoe Jing Xu & Alejandro Lleras

Traditionally, researchers interpret the difficulty in conjunction search as difficulty in binding features. In the present study, we used a behavioral-computational approach to assess if parameters from feature search could predict performance in a color-shape conjunction search task. We also investigated whether pooling-mediated processing in peripheral regions was a key-limiting factor in performance in conjunction search by manipulating display arrangements across different experiments. The results indicated that parameters in homogeneous search displays can indeed be used to successfully predict performance in conjunction search displays. This finding is noteworthy because it indicates that the visual system must be extracting the same information from the display in feature and conjunction search tasks (i.e., the target-distractor similarity) using color and shape. Furthermore, there was no compelling evidence that pooling-mediated processing was the primary constraint on performance in this conjunction search task. A model-comparison approach compared the accuracy of different distractor rejection architectures in predicting performance in conjunction search tasks. The winning model showed participants engaging hierarchically with the display, selecting and rejecting distractor subsets based on a single defining feature. Taken in the context of previous research on heterogeneous search performance, the current results imply that the inherent demands of search for a conjunction of color and shape compel participants to adopt this targeted search strategy.

Some visual searches are fast and unfold in parallel in circumstances where the target features sufficiently differ from those of the distractors, such that peripheral vision can discriminate between the two. This occurs in feature search, where the target has a unique feature that is not shared by any of the distractors. In this condition, Treisman and Gelade[1] found that the reaction times were independent of set size, which was indicative that the target had been detected in parallel. The authors proposed that parallel detection of the target was possible because there existed one feature map that was activated by a single item in the display (the target).

Other search conditions have longer response times and cannot unfold in parallel. This is the case for conjunction search, where the target is accompanied by distractors of different types, with each type sharing one feature with the target. Treisman and Gelade[1] showed that, in conjunction search, search performance suffers compared to feature search, with response times increasing substantially, as a linear function of set size. Feature Integration Theory[1] proposed that parallel detection was not possible because all active feature maps were coding multiple items at a time. Instead, spatial attention was necessary to bind all features present at a given location into coherent object representations that could then be compared to the target template. This serial comparison process was responsible for the steep and linear search functions.

Since, several studies have questioned the seriality argument in conjunction search. For example, Eckstein and colleagues[2,3] found that an unlimited-capacity signal detection theory-based model can account for conjunction search data without requiring visual attention for binding the feature information serially. Further, by comparing simultaneous and successive presentation of subsets of the display, Huang and Pashler[4] found that there were little attentional capacity limitations in conjunction search. Finally, recent computational modeling work has shown that focused spatial attention is not required to bind two features (like color and shape) into object representations, and in fact, search for targets defined by two features can unfold in parallel and be guided by information along the two feature dimensions (e.g., Buetti, Xu & Lleras[5]; Xu, Lleras & Buetti[6]).

More recently, Rosenholtz and colleagues[7,8] put forward an account of search that does not rely on spatial attention to serially move between locations to explain the difficulty in conjunction search, but rather focuses on the essential role peripheral vision plays in processing visual scenes. Their model (i.e., the Texture Tiling Model) is based on the idea that representations of visual information in the periphery are the result of summary statistics computed over local pooling regions, which increase in size with increasing eccentricity. These

Department of Psychology, University of Illinois Urbana-Champaign, Champaign 61820, United States. email: yaoyunc2@illinois.edu

summary statistics are the same ones that are believed to be responsible for texture perception[9] and capture quite effectively the phenomenon of visual crowding. Extending this approach to visual search, the authors propose that performance in visual search is best understood when one considers the peripheral discriminability of patches of the image containing both target and distractors and patches only containing distractors, in terms of their summary statistics. Indeed, search speeds were shown to decrease as the discriminability between these two types of patches decreased. In the case of conjunction search, the Texture Tiling Model can easily account for the difficulty observed in the typical conjunction search task. This follows because pooling regions containing two different types of distractors are likely to have summary statistics comparable to those of a pooling region containing the target and some distractors.

Using mathematical modeling to uncover underlying processing

Mathematical modeling can be used to differentiate performance across different cognitive architectures[10–12]. In particular, when the processing is assumed to be parallel in nature, there are several possible cognitive architectures that can produce efficient search behavior, yet each architecture has its own mathematical signature. Thus, these formulae can predict search performance and identify the architecture that best fits observed behavior. For example, parameters from homogeneous searches can successfully predict performance in heterogeneous searches when target and distractors are sufficiently different such that peripheral vision can distinguish between them[12–14]. Figure 1 shows studies using this behavioral-computational approach.

The general idea of this approach is to first measure search slopes under parallel, distractor homogeneous search conditions, using those as an index of the visual system's ability to detect the target among identical distractors. Next, search times are measured when the target appears in distractor heterogeneous displays. Finally, one posits different architectures regarding how distractors are rejected in the heterogeneous condition, and uses parameters from homogeneous conditions to predict search times in heterogeneous conditions. For all scenarios in Figure 1, the winning model is the Parallel-Simultaneous Rejection model (see methods section for a description of all models), which assumes all distractor types are processed simultaneously from the start of the trial. Each distractor is then rejected at its own rate, determined by its similarity to the target, as indexed by the search slope observed in homogenous displays. Note that in this section we are referring to search slopes that are computed in a logarithmic scale, as these logarithmic search slopes provide a better account of homogeneous search performance than linear search functions (e.g., Buetti et al.[11]; Lleras, et al.[15]). These studies found no qualitative processing difference between homogeneous and heterogeneous searches: when different types of distractors are spatially intermixed, each distractor type is simply rejected at a multiplicatively slower rate than when only one type of distractor is presented on the display. Finally, it should be noted that the studies in Figure 1 used targets and distractors that shared no features, unlike a typical conjunction search. Consequently, whether the parameters from parallel, feature search scenarios can predict performance in conjunction search conditions remains an open question.

Present study

The present study aimed to assess the extent to which conjunction search performance can be understood based on feature search performance. We adopted the behavioral-computational approach illustrated in Figure 1 to examine distractor rejection in conjunction search task. We began by estimating the logarithmic search slopes observed in feature search tasks, using the same visual features that will be used to construct the conjunction search task: in one condition, we will estimate the efficiency at which observers can detect a target defined by the same color difference used in the conjunction search (here, finding a red target among orange distractors), as well as the efficiency at which they detect the shape target (a triangle target among circle distractors). These slopes were then used to fit parallel processing models to predict response times in the corresponding color-shape conjunction search task.

If the same underlying feature comparisons that govern processing times in feature search are also at play in the color-shape conjunction search, then our model-and-predict approach should be able to capture a vast majority of the variance in conjunction search performance. However, if the color-shape conjunction search fundamentally differs from feature search, such as by relying on serial scrutiny of stimuli, or relying on a different type of comparison (foveal scrutiny of stimuli, rather than peripheral, parallel comparisons), then the parameters derived from parallel processing in feature search may be less effective in predicting the color-shape conjunction search performance.

In addition, we also manipulated display configurations to evaluate the role of pooling-mediated processing in color-shape conjunction search difficulty. We tested whether pooling-mediated processing induces the difficulty in conjunction search by comparing performance between different types of display arrangements - some provided opportunities for feature confusability inside pooling regions (traditional, rectangular grid displays) while others minimized the confusability (circular grid, spaced elements beyond Bouma's law, and cortically-magnified stimuli). If color-shape conjunction search difficulty is caused by feature confusability in peripheral vision for these features, then spacing stimuli should significantly facilitate search performance, perhaps to the point where performance is qualitatively similar to the one observed in feature search conditions.

To further reduce confusability within pooling regions, we spatially segregated distractors by type in two experiments, which ensured local homogeneity among the items. Based on previous findings, we expected spatial segregation to accelerate the rate of distractor processing[13], possibly matching feature search conditions. Indeed, Lleras et al.[13] studied performance in search tasks with heterogeneous search displays (i.e., displays containing two different types of distractors, *D1* and *D2*, that do not share any features with the target, *T*, unlike conjunction search) and demonstrated that (i) RTs in heterogeneous conditions were predicted by the logarithmic slopes from corresponding homogeneous searches (search for *T* among *D1*, and search for *T* among *D2*); (ii) when distractors were spatially intermixed, processing rates were multiplicatively longer, slowing down

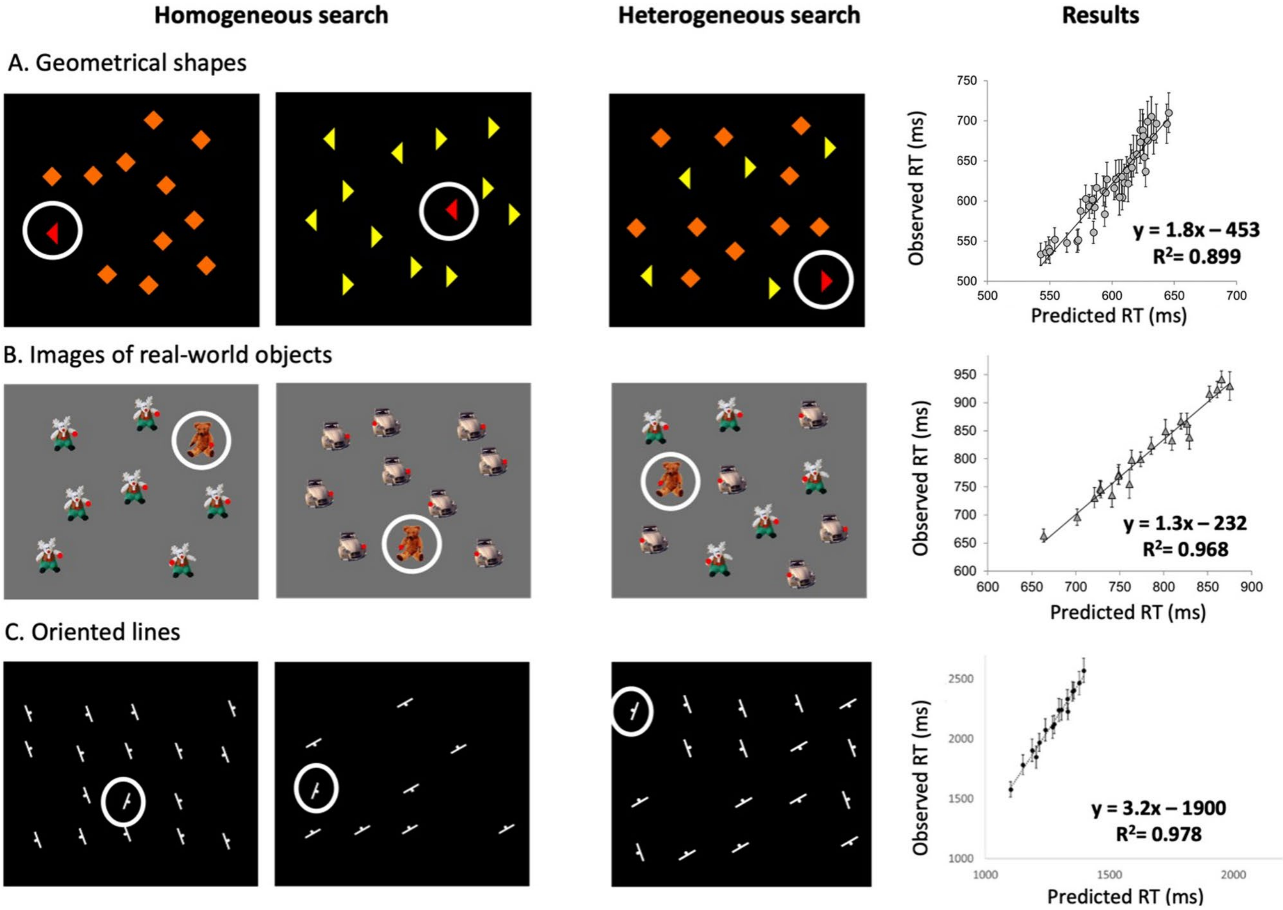

Fig. 1. Illustration of the approach. Examples of the experimental procedure used in three previous studies. The two displays on the left side of the figure illustrate homogeneous search conditions that are used to estimate the parameters (logarithmic search slopes) that will be used to predict performance in heterogeneous search displays that use the same target and distractor stimuli. The heterogeneous displays are illustrated in the middle panel. In A and B, heterogeneous displays could contain two or three different types of distractors on any given display, and the set size of each type of distractor was varied independently. The white circle was not present in the experiments, it is used here to highlight the target. On the right, the result figures illustrate the success of the winning model (see text for more details) by plotting the predicted RT expected from the model against the observed RT obtained in the heterogeneous experiment. The R^2 value illustrates the amount of variance accounted for by the winning model. The slopes of the fitted regression lines are all larger than 1, indicating that distractor rejection times were multiplicative slowed down in the heterogeneous conditions. When distractors are spatially segregated by type (distractor of type one on the left and distractor of type two on the right), the slope is close to one (not shown). Finally, note that different participants completed the homogeneous and heterogeneous search conditions. Data from A, B, and C come from Lleras et al.[13], Wang et al.[12] and Xu et al.[14], respectively.

search performance, compared to homogeneous searches; (iii) when distractors were spatially segregated (e.g., *D1* distractors on the left, *D2* distractors on the right), processing rates were comparable to those observed in homogeneous searches. This latter finding was interpreted as reflecting inter-item interactions: when nearby-items are similar to one another, they facilitate each other's processing (i.e., rejection becomes faster - note that Lleras et al.[15]suggested that multiplicative changes in processing rates can also be interpreted as changes in the noise present during the evidence accumulation process. When all items near one another are identical, all the evidence accumulated is of a same kind, whereas when distractors are intermixed, evidence accumulation at any one location can become noisier by the presence of different sensory signals nearby.) Duncan and Humphreys[16] had referred to a similar mechanism as spreading suppression, which is a form of perceptual grouping. Therefore, we should observe easier searches in spatially segregated displays than intermixed displays.

Experiment 1: Feature search

The search displays were either rectangular grids (Experiment 1A) or concentric circles with cortical magnification compensation (Experiment 1B). The slope (*D*) values for the color and shape search efficiencies observed in this experiment were used to predict search performance in the the color-shape conjunction search conditions studied in Experiment 2.

Methods

The methods and experimental protocols of this experiment (and later experiments) were approved by the Institutional Review Board at the University of Illinois Urbana-Champaign, and are in accordance with the Declaration of Helsinki.

Participants

Participants for this set of experiments were recruited from the University of Illinois Urbana Champaign. They received course credit in exchange for compensation. Note that due to COVID-19 pandemic, all experiments were run online using participants' own devices. Only participants who self-reported to have normal or corrected-to-normal vision and normal color vision were allowed to take part in the experiment. All study participants provided informed consent.

Past experiments in our lab have showed that 35 participants were adequate for successful predictive analyses using online data collection[14,17]. Due to unexpectedly high signup rates near the end of the semester, we ended up collecting fifty-three participants in Experiment 1A (15 males, 38 females; mean age = 19.27; age range: 18 - 22), and forty-seven participants in Experiment 1B (11 males, 36 females; mean age = 19.13; age range: 18 - 22) before we halted data collection.

The criteria for data inclusion are based on accuracy and reaction time analyses. The accuracy rate was calculated by dividing the number of correct trials by the number of trials the participants responded to (that is, time-out trials were not included in the computation of accuracy as they could result from instability in internet connection). The following trials were not included in the analysis: trials in which the participants pressed the wrong button (incorrect trials), and trials in which the participants provided no response after 2.5 seconds (time-out trials). Furthermore, participants' data were excluded if their time-out rate was greater than 15% or if their overall accuracy (including incorrect trials but not time-out trials) was less than 90%. Finally, participants whose reaction times were more than 2.5 standard deviations away from the mean reaction time across all participants were also excluded.

Twelve participants were excluded from Experiment 1A (six of them were excluded because of low accuracy and six because of high time-out rate). Ten participants were excluded from Experiment 1B (eight of them were excluded because of low accuracy and two because of high time-out rate). The data analysis included 41 participants in Experiment 1A (mean accuracy = .97, accuracy SD = .024; after excluding incorrect trials, mean response time = 813 ms, response time SD = 67 ms), and 37 participants in Experiment 1B (mean accuracy = .97, accuracy SD = .023; after excluding incorrect trials, mean response time = 892 ms, response time SD = 152 ms)

Stimuli and design

Due to COVID-19, participants completed the experiments online using their own devices. All experiments were programmed in JavaScript and conducted on Pavlovia (pavlovia.org). Because the experiments were run online and we had no control over the visual angle of the stimuli on participants' computers, at the beginning of the experiment, we asked the participants to rescale an image of a credit card to match its size to a real one, ensuring that the stimuli across different equipment had the same physical size.

Experiment 1A: Rectangular Grid Search Display. The stimuli in Experiment 1A were randomly placed on the display based on an invisible 6 by 6 grid occupying an area of 25 cm × 14 cm on the center of participants' screens, with a total of 36 possible locations. This size was chosen to allow the participants with a screen as small as 12.5 inches to see the full display clearly. Jitter of up to 25 pixels allowed the position of the stimuli to vary slightly inside each grid position.

The target was a red triangle with a white dot on either its left or right. There were no target-absent trials. The displays in Experiment 1A were always homogeneous, i.e., the distractors on display were always of the same type (except in the target-only condition, there was no distractor). The task was to respond by pressing a key to indicate which side the white dot was on. There are two different types of search tasks, which were blocked and counterbalanced. In the color search block, distractors were orange triangles. In the shape search block, distractors were red circles. Participants were randomly assigned to begin with either color search or shape search (sample displays shown in Figure 2).

There were five possible set sizes for the distractors: 0, 2, 4, 8, and 16. In total, there were 20 conditions (search type by set size by target white dot position) that were repeated 35 times for a total of 700 trials.

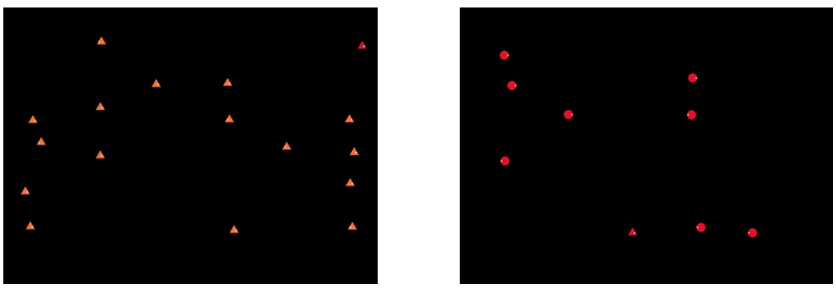

Fig. 2. Examples of search displays in Experiment 1A. Left: Sample display of color search, with orange triangles (distractor type 1) as distractors. Right: Sample display of shape search, with red circles (distractor type 2) as distractors.

Experiment 1B: Concentric Search Grids. The stimuli in Experiment 1B were randomly placed in the display based on three invisible concentric circles (centered on the center of the display), each with 12 evenly spaced possible locations (36 in total). The eccentricities of these three circles were 4.5°, 8.4°, and 15.4°of visual angle. The search items were positioned in this way so that there was symmetry in the display along the vertical midline. According to Bouma's equation[18], this concentric separation can minimize crowding effects. The concentric circles had different magnification multipliers (M) as cortical compensation which were adapted from Wang et al.[19], following Rovamo and Virsu[20], where M was an approximately linear function of eccentricity. Random jitters within 25 pixels allowed the position of the stimuli to vary slightly.

The goal of these homogeneous search displays in concentric circles with cortical magnification compensation was to find slope D values for color search and shape search under conditions that minimize the likelihood that two or more stimuli would be processed inside the same peripheral pooling region (see Rosenholtz, Huang, Raj, et al.[8]).

The target, distractors, and blocks settings were identical to those in Experiment 1A except that the possible set sizes for the distractors were 0, 2, 4, 8, 16, and 32, and that there was a white fixation in the center along with the search items to help participants fixate the center of the display throughout the trials. In total, there were 24 conditions that were repeated 35 times for a total of 840 trials (sample displays shown in Figure 3).

Procedure

In Experiment 1A and 1B, after giving consent and reading the instructions, the participants completed ten practice trials. At the beginning of each trial, participants were shown a white fixation cross at the center of the screen, followed by the search display. Participants were asked to search for the target and report the left or right location of the white dot by pressing the left or right arrow key on their keyboard. Each display lasted for a fixed amount of time (2.5 seconds in Experiment 1A; 5 seconds in Experiment 1B) or until the participants pressed their response key, whichever occurred earlier. After each response, a visual feedback of "Correct!" or "Wrong!" was displayed, lasting for 0.5 second. The trial ended with a black background shown for 0.5 second. The participants were given an optional break every 75 trials throughout the whole experiment.

Results and discussion

In Experiment 1A (rectangular grid), the logarithmic slopes for color search and shape search were 62.1 and 95.0, respectively (unit: ms/log unit of set size, see Figure 4). In Experiment 1B (circular grid), the logarithmic slopes for color search and shape search were 30.4 and 79.5, respectively (unit: ms/log unit of set size). The search functions in Experiment 1 were all well fitted by a logarithmic function (R^2between 0.96 and 0.99). As expected, the results showed that the circular grids produced relatively smaller slopes in both color search and shape search compared to the rectangular grids[19].

Experiment 2: Color-Shape conjunction search

The search times from experiments where distractors were intermixed (2A) or spatially segregated (2C) on a rectangular grid were predicted by using the search slopes from Experiment 1A. The search times from experiments where distractors were intermixed (2B) or spatially segregated (2D) on a circular grid were predicted by using the search slopes from Experiment 1B.

Methods

Participants

The participant recruitment procedure was identical to that in Experiment 1. There were fifty-three participants in Experiment 2A (13 males, 40 females; mean age = 19.31, age range: 18 - 22), forty-five participants in Experiment 2B (16 males, 29 females; mean age = 19.11, age range: 18 - 25), forty-two participants in Experiment 2C (13 males, 29 females; mean age = 18.56, age range: 18 - 22), and fifty-three participants in Experiment 2D (16 males, 36 females, 1 other; mean age = 18.98, age range: 18 - 21). All study participants provided informed consent.

The criteria for data inclusion were the same as that in Experiment 1. Eight participants were excluded from Experiment 2A (seven of them were excluded because of low accuracy and one was excluded because of high time-out rate). Eleven participants were excluded from Experiment 2B (seven of them were excluded because of

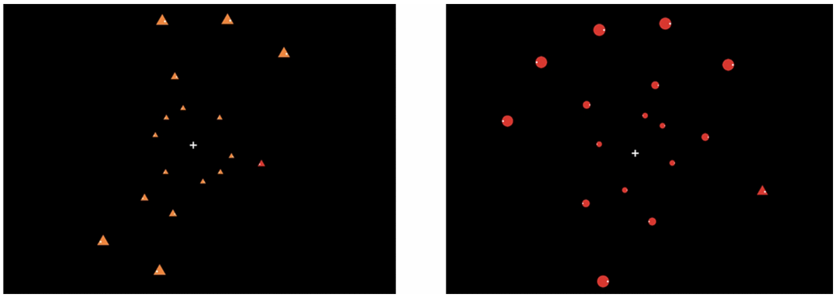

Fig. 3. Examples of search displays in Experiment 1B. Left: Sample display of color search, with orange triangles (distractor type 1) as distractors. Right: Sample display of shape search, with red circles (distractor type 2) as distractors.

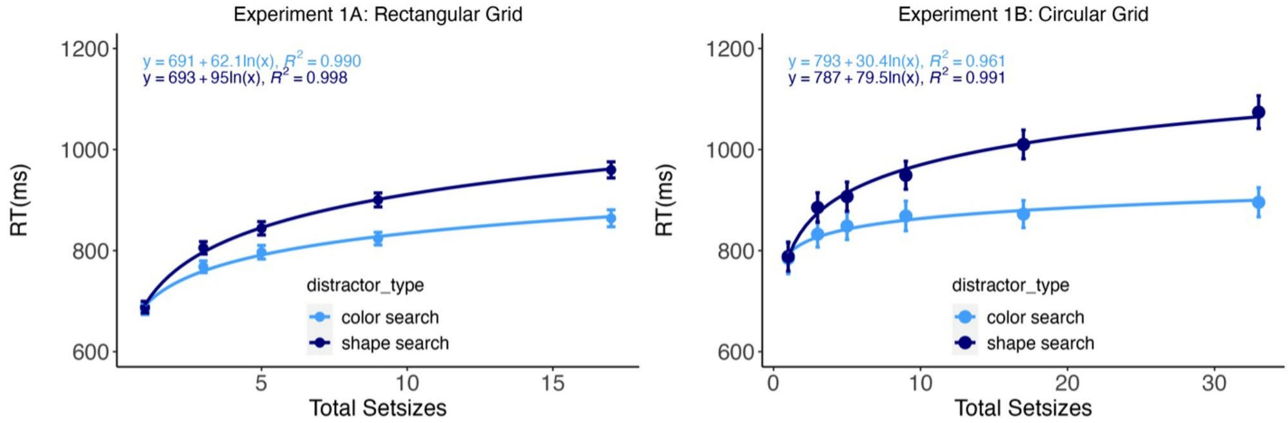

Fig. 4. Results of Experiment 1 with rectangular grids and circular grids. Response times observed in the homogeneous search displays were plotted as a function of total set sizes and distractor types. Dotted lines showed the best logarithmic fit for the data. Error bars showed one standard error of the mean. Refer to the online article for the color version of this figure.

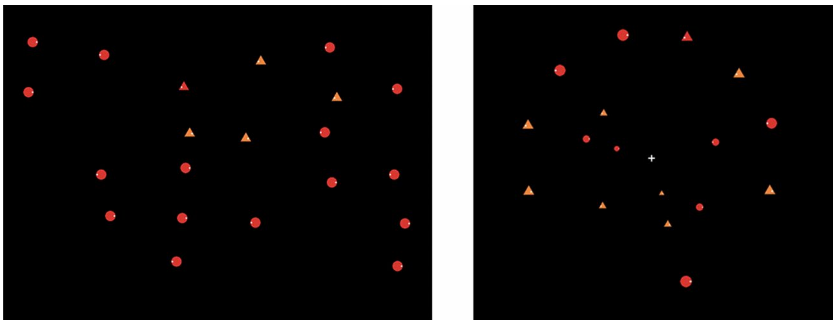

Fig. 5. Example of search displays in Experiment 2A and 2B. The left image shows an example search display in Experiment 2A, and the right image shows an example search display in Experiment 2B.

low accuracy, three were excluded because of high time-out rate, and one was excluded because of their RT being an outlier). Eight participants were excluded from Experiment 2C (seven of them were excluded because of low accuracy, and one was excluded because of their RT being an outlier). Fourteen participants were excluded from Experiment 2D (because of low accuracy).

The data analysis included 45 participants in Experiment 2A (mean accuracy = .97, accuracy SD = .019; after excluding incorrect trials, group mean response time = 1226ms, response time SD = 171ms), 34 participants in Experiment 2B (group accuracy = .97, accuracy SD = .020; after excluding incorrect trials, group mean response time = 1118ms, response time SD = 180ms), 34 participants in Experiment 2C (group accuracy = .98, accuracy SD = .020; after excluding incorrect trials, group mean response time = 1159ms, response time SD = 169ms), and 39 participants in Experiment 2D (group accuracy = .97, accuracy SD = .020; after excluding incorrect trials, group mean response time = 1119 ms, response time SD = 216 ms).

Stimuli and design

It is worth noting that traditional conjunction search studies typically do not independently vary set size for the different types of distractors, but rather use a similar amount of distractors of each kind on every display[1,21–25]. However, in the current experiments, following the methodology used in Xu et al.[14], we used conditions where we independently manipulated the number of distractors from each type.

Experiment 2A: Intermixed Distractors in Rectangular Grid Displays. The stimuli and settings in Experiment 2A were the same as in Experiment 1A except that both types of distractors (orange triangles and red circles) were simultaneously displayed in each trial (except for the target-only condition where there was no distractor).

The displays in Experiment 2A were color-shape conjunction search displays (except for the target-only condition), which means both types of distractors were present in the display (a sample display is shown in Figure 5 left panel). Set size was manipulated independently for each distractor type (2, 4, 8, 16 distractors). This resulted in 16 unique conditions, plus a target-only condition. With the two dot placements on the target (left, right), that amounts to a total of 34 conditions. Each condition was tested 20 times, resulting in 680 trials overall.

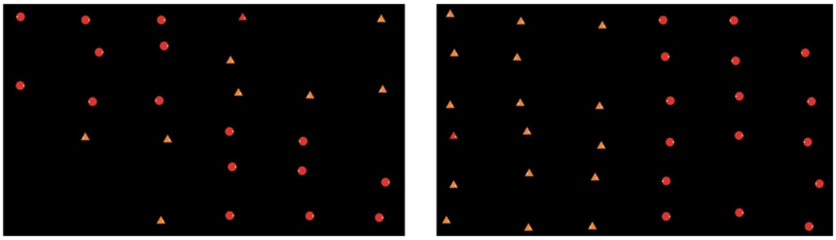

Fig. 6. Example of search displays in Experiment 2C.

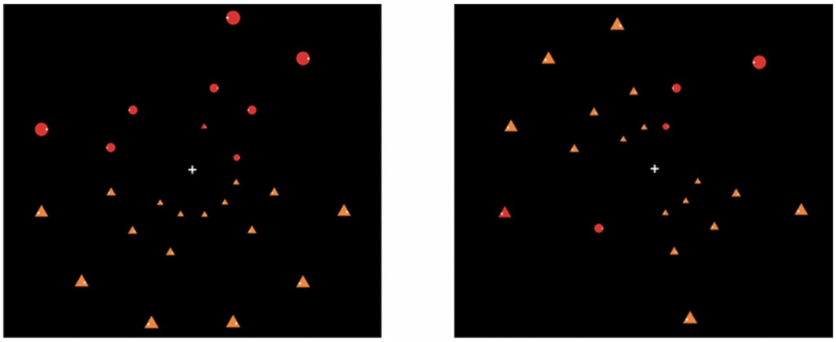

Fig. 7. Example of search displays in Experiment 2D.

Experiment 2B: Intermixed Distractors in Circular Grid Displays. The stimuli and display configuration were the same as in Experiment 1B. The experimental design was identical to Experiment 2A: same set sizes and distractor combinations, and same number of observations (Figure 5, right). As a reminder, the display configuration aimed at minimizing the chances that two different stimuli would be processed by the same peripheral pooling region.

Experiment 2C: Segregated Distractors in Rectangular Grid Displays. Everything in Experiment 2C was identical to Experiment 2A, except for the following. In Experiment 2C, the distractors were arranged in local homogeneous regions. Specifically, the rectangular grid was divided into four quadrants (top left, top right, bottom left, bottom right, see Figure 6) and each quadrant of the display contained only one type of distractor.

Experiment 2D: Segregated Distractors in Circular Grid Displays. Everything in Experiment 2D was identical to Experiment 2B, except for the following. In Experiment 2D, the distractors were arranged in local homogeneous regions. As in Experiment 2C, the circular grid was divided into four quadrants (top left, top right, bottom left, bottom right, see Figure 7) and each quadrant of the display contained only one type of distractor.

Procedure

In Experiment 2, the procedure was identical to that in Experiment 1B and the participants were given an optional break every 175 trials (Experiment 2A) and every 68 trials (Experiment 2B, 2C, and 2D) throughout the whole experiment. Within each sub-experiment, all conditions were intermixed. In terms of predicting RTs, for each experiment, the D_j parameters were extracted from the corresponding type of grid in homogeneous conditions: for predicting performance in Experiments 2A and 2C, we used D_j parameters from Experiment 1A, because the same rectangular grid display was used across these three experiments, whereas for predicting performance in Experiment 2B and 2D, the D_j parameters from Experiment 1B were used (all circular grid displays).

Analyses

In all analyses, the response time analyses only included correct trials. The datasets generated and/or analysed during the current study are available in the Open Science Framework repository, link: https://osf.io/ymv28/.

Candidate models for model comparison

Wang et al.[12] developed different computational models to capture different parallel processing architectures.

Model 1: ***Parallel-Simultaneous Rejection Model***. This model assumes that in heterogeneous displays, distractors of type 1 and 2 are processed simultaneously from the start of the trial, with each distractor being individually rejected at its own rate. The rejection rate is set to be the same as the one observed in homogeneous search conditions. Equation (1) presents the model for the case where there are two types of distractors in the display.

$$RT = a + D_1 \times \ln(N1 + N2 + 1) + (D_2 - D_1) \times \ln(N_2 + 1) \quad (1)$$

Where the constant *a* represents the reaction time when the target is alone in the display. N_1 and N_2 represent the number of distractors of type 1 and 2 in the display, and D_1 and D_2 denote the logarithmic slope parameters associated with distractors of type 1 and 2 (organized from smallest D_1 to largest D_2), estimated using homogeneous displays.

Model 2: *Parallel-Sequential Rejection Model.* This model assumes that one of the two sets of distractors is attended to and rejected in parallel before the second set is attended to and rejected in parallel, as shown in Equation (2). In this model, there are two types of distractors, and although the model does not allow us to conclude which distractor type will be processed first, if this model wins, it means that the two groups of distractors are processed in turn. Critically, the difference between Model 2 and Model 1 lies in the fact that Parallel-Sequential Rejection Model does not allow simultaneous processing of both types of distractors but focuses processing only on one type of distractor at a time. This would suggest that there is a cognitive strategy component in conjunction search tasks (e.g., first discard D_1 distractors, then discard D_2 distractors).

$$RT = a + D_1 \times \ln(N_1 + 1) + D_2 \times \ln(N_2 + 1) \tag{2}$$

The Parallel-Sequential Rejection Model is consistent with previous findings[26–29], if one assumes that the first set of distractors to be attended-and-rejected are those defined by color. Model 2 was found to be the winning model in more recent work that evaluated the rejection process when stimuli are defined by color-color conjunctions (e.g., find a red-center/green-surround target among green-center/red surround distractors, Wang et al.[30]; Wang et al.[31]).

Model 3: *Single-Threshold Model.* This model assumes a single decision criterion can be used to distinguish targets from distractors, indexed by the largest of the slopes D_1 and D_2, as shown in Equation (3). To be specific, the activation signals start to accumulate in parallel for items across all locations and the target-distractor differentiation is a result of the activation map calculated with a single decision threshold after the accumulation stage ends. With the identical time constant cost parameter D_{max} for all items across the display during the evidence accumulation processes, this model implies that all distractors are processed at one fixed speed depending on the overall contrast signal between distractors and the target (unlike Model 1, it does not have varying processing rates). More importantly, this model stands in for all the models that treat all distractors in the same way, that is, with a single decision threshold (e.g., Zelinsky, 2008[32]; Wolfe, 1994[33]).

$$RT = a + D_{\max} \times \ln(N_T + 1) \tag{3}$$

Here, N_T represents the total number of distractors on the display.

Finally, to evaluate the performance of each model, predicted RTs by each model were regressed against observed RTs in conjunction search experiment, as indicated by Equation (4). The R^2 value represents the degree of variability in the observed data that can be predicted by each model. The β parameter provides an indication of how processing times in heterogeneous conditions compare to processing times in homogeneous conditions. When $\beta \approx 1$, it indicates the time taken to reject a distractor in the heterogeneous display is similar to the time taken to reject the same distractor in a homogeneous display; when $\beta > 1$, it indicates the rejection times for individual distractors are multiplicatively slower (by a factor of β) in heterogeneous conditions compared to homogeneous displays.

$$\mathrm{RT}_{\text{observed}} = C + \beta \times \mathrm{RT}_{\text{predicted}} \tag{4}$$

In Equation (4), C and β are the free parameters estimated based on Ordinary Least Squares. Refer to Figure 8 for the graphical representation of this linear regression.

Model comparison approach

We used the Akaike Information Criterion (AIC) model comparison metric to do model evaluation based on the likelihood of the different models. We evaluated the strength of evidence in favor of the winning model by computing $\exp((\mathrm{AIC}_{\min} - \mathrm{AIC}_i)/2)$.

Bootstrapping analysis

The free parameter β in Equation (4) was proposed to capture the slowing down effect caused by intermixing distractors[13]. To ascertain whether the β values calculated based on Equation (4) were significantly greater than 1, we conducted a bootstrapping analysis. In this analysis, for each experiment, we resampled the participants' trial data with replacements in the experiments. We maintained a sample size of 40 participants in each iteration and repeated this process 100 times to obtain a distribution of β values, resulting in 100 β values for the winning model. We then conducted a one-sample t-test to evaluate whether the β value was significantly larger than 1.

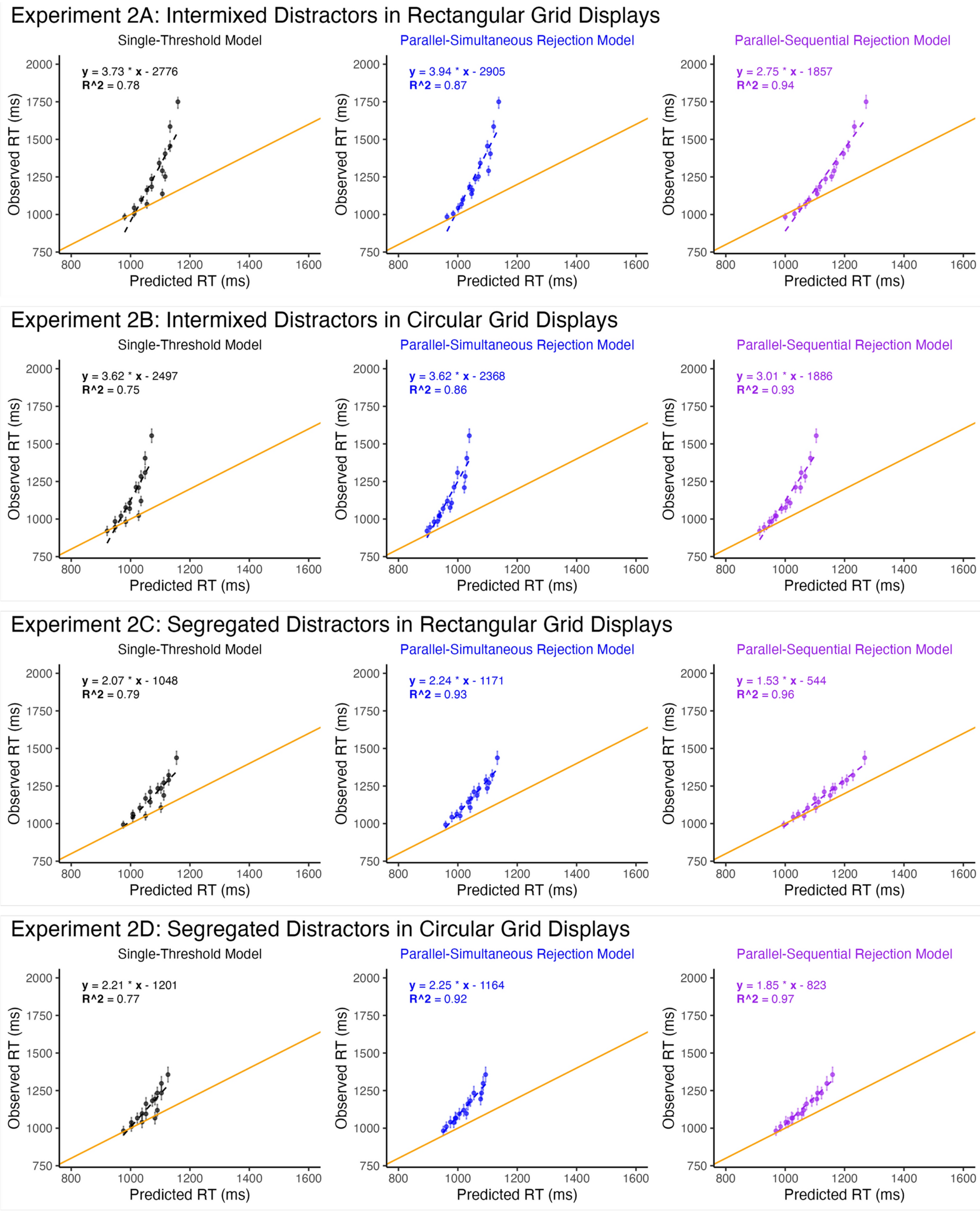

Fig. 8. Model performance plots. Observed reaction times plotted against the predicted reaction times in Experiments 2A (the top row), 2B (the second row), 2C (the third row), and 2D (the fourth row) for Single-Threshold Model (left), Parallel-Simultaneous Rejection Model (middle), and Parallel-Sequential Rejection Model (right), respectively. The dotted lines showed the best linear fit for the data. Error bars showed one standard error of the mean. The orange lines refer to y = x.

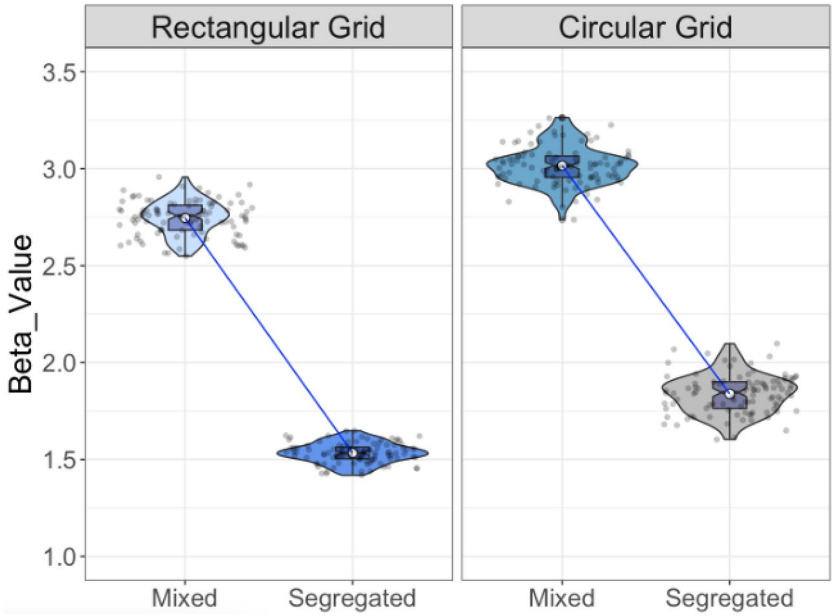

Fig. 9. Distributions of β values from bootstrapping analysis. The beta values of the Parallel Sequential Rejection Model from the bootstrapping analysis, presented in violin plots. Each violin plot displays the distribution of β values across all iterations within a single experiment - specifically, the Mixed condition under the Rectangular Grid refers to Experiment 2A, the Segregated condition under the Rectangular Grid refers to Experiment 2B, the Mixed condition under the Circular Grid refers to Experiment 2C, and the Segregated condition under the Circular Grid refers to Experiment 2D. Accompanying boxplots illustrate the interquartile range (boxes), the median (horizontal notched markers), and the mean (white circles connected by blue lines). Light gray circles represent individual β values from each of the iterations of the simulation, while the shaded regions indicate the data's probability density.

Between-experiment comparison on β values.
To evaluate the efficiency of distractor rejection across different spatial arrangements, as indexed by β, we utilized results from bootstrapping analyses. We conducted Welch two-sample t-tests, assuming unequal variances, to compare the β values between segregated and intermixed spatial arrangements: Experiment 2A versus 2C (both featuring rectangular grid displays) and Experiment 2B versus 2D (both with circular grid displays).

Results and discussion
Experiment 2A: Intermixed distractors in rectangular grid displays
The observed response times for all distractor combination conditions in Experiment 2A-D are shown in the appendix (Table S1-S4). Among the three models, the R^2 associated with the Parallel-Sequential Model ($R^2 = 0.94$) was always larger than the R^2 associated with the Parallel-Simultaneous Rejection Model ($R^2 = 0.87$) and Single Threshold Model ($R^2 = 0.78$). The Akaike Information Criterion (AIC) model comparison showed that the Parallel-Sequential Rejection Model (AIC = 175.91) was substantially more likely than the other two models: it was 5.88×10^4 times more likely than the Single Threshold Model (AIC = 197.88), and 744 times more likely than the Parallel-Simultaneous Rejection Model (AIC = 189.14).

The bootstrapping analysis (distribution illustrated in Figure 9) showed a mean β of 2.75, which was significantly larger than 1, $t(99) = 197.78$, $p < 0.0001$, 95% CI: [2.73, Inf]. This indicated a slow-down in rejecting individual distractors when spatially intermixing distractors compared to when distractors are all homogeneous.

Experiment 2B: Intermixed distractors in circular grid displays
The R^2 associated with Parallel-Sequential Model ($R^2 = 0.93$) was larger than the R^2 associated with Parallel-Simultaneous Rejection Model ($R^2 = 0.86$) and Single Threshold Model ($R^2 = 0.75$). The Parallel-Sequential Rejection Model (AIC = 174.23) was 1.88×10^4 times more likely than the Single Threshold Model (AIC = 193.91) and 174 times more likely than the Parallel-Simultaneous Rejection Model (AIC = 184.55).

The bootstrapping analysis showed a mean β value of 3.02, significantly larger than 1, $t(99) = 203.97$, $p < 0.0001$, 95% CI: [3.00, Inf]. This indicated that spatially intermixing distractors slows down search compared to when displays are homogeneous, even when stimuli are arranged to minimize the possibility of crowding.

Experiment 2C: Segregated distractors in rectangular grid displays
The R^2 for the Parallel-Sequential Rejection Model ($R^2 = 0.96$) was again higher than the R^2 associated with the Parallel-Simultaneous Rejection Model ($R^2 = 0.93$) and Single Threshold Model ($R^2 = 0.79$). The Parallel-

Sequential Rejection Model (AIC = 150.50) was 8.92×10^5 times more likely than the Single Threshold Model (AIC = 177.91) and 217 times more likely than the Parallel-Simultaneous Rejection Model (AIC = 161.26).

The bootstrapping analysis showed a mean β value of 1.53, again significantly larger than 1, $t(99) = 102.93$, $p < 0.0001$, 95% CI: [1.53, Inf] (see Figure 9). This indicated that although distractors were spatially segregated, participants still performed the search slower than under the comparable entirely homogeneous conditions of Experiment 1A. The bootstrapping analysis also indicated that the segregation manipulation, nonetheless, significantly accelerated the rejection of distractors by a factor of 1.8 (i.e., 2.75/1.53), compared to when distractors were intermixed (in Experiment 2A, β was 2.75). This improvement in performance is consistent with what was proposed by Lleras et al.[13], that when nearby distractors are identical to one another, they facilitate rejection of one another in a multiplicative way. The fact that β was larger than 1 might simply reflect the fact that rejection unfolded in a sequential manner. Instead of processing all items in the display at the same time, participants reject distractors sequentially, by quadrant, or by subset or both.

Experiment 2D: Segregated distractors in circular grid displays

The R^2 for the Parallel-Sequential Rejection Model (0.97) was once again the highest among the three models (Single Threshold Model $R^2 = 0.77$, and Parallel-Simultaneous Rejection Model $R^2 = 0.92$). The Parallel-Sequential Rejection Model (AIC = 142.96) was 1.50×10^7 times more likely than the Single Threshold Model (AIC = 176.00) and 3308 times more likely than the Parallel-Simultaneous Rejection Model (AIC = 159.17).

The values of β for the winning model was smaller (1.85) than when distractors were intermixed (3.01 in Experiment 2B), as confirmed by the bootstrapping analysis (Figure 9) – a mean β value of 1.84, significantly larger than 1, $t(99) = 84.82$, $p < 0.0001$, 95% CI: [1.82, Inf]. In other words, the efficiency of distractor rejection processes was slower by a factor of 1.85 when displays were spatially segregated, compared to when they were homogeneous, meaning that, unlike Lleras et al.[13], distractor segregation did not reinstate the ease to reject distractors that is observed under homogeneous conditions (Experiment 1B). That said, beta was smaller than in Experiment 2B: distractor rejection was accelerated by a factor of 1.6 (i.e., 3.01/1.85) times by the spatial segregation manipulation.

The findings from this study demonstrate that spatial segregation of distractors improves search performance, as reflected by the reduced β values in Experiments 2C and 2D compared to their intermixed counterparts (Experiments 2A and 2B). In Experiments 2C and 2D, configurations where two adjacent quadrants contained the same type of distractors (i.e., half of the display shared the same distractor type) yielded β values closer to 1 (1.3 in Experiment 2C, 1.62 in Experiment 2D), compared to configurations where adjacent quadrants contained different distractor types (1.93 in Experiment 2C, 2.31 in Experiment 2D). These results suggest that pooling-mediated processing plays a role in search efficiency, particularly when distractor types are intermixed. However, the persistence of β values greater than 1 under optimal conditions (e.g., spatial segregation) indicates that pooling alone does not fully account for the slowdown observed in color-shape conjunction search tasks. Instead, the Parallel-Sequential Rejection Model highlights the importance of hierarchical processing strategies inherent to the color-shape conjunction search task in determining search performance.

Between-experiment comparison on β values

The results indicated significant differences in β values under both types of display conditions. For the rectangular grids, the β values differed significantly between Experiment 2A and 2C, $t(160.03) = 118.42, p < 0.0001$, 95% CI: [1.19, 1.23]. Similarly, for the circular grids, there was a significant difference between Experiments 2B and 2D, $t(198) = 84.21$, $p < 0.0001$, 95% CI: [1.15, 1.20]. These results suggested that segregating distractors can indeed accelerate distractor rejection, consistent with what Lleras et al.[13] proposed.

General discussion

The goal of the present study was to try to determine the factors responsible for the slowdown in search performance observed in color-shape conjunction search compared to the performance observed in the corresponding feature-search tasks that use the same visual discriminations. The results in Experiment 2 showed that the behavioral-computational approach[12–14]successfully shed light on the underlying mechanisms behind distractor rejection processes in a color-shape conjunction search: the finding that the parameters governing homogeneous feature search performance continue to be the main determinants of performance under conjunction search conditions implies that the visual system consistently extracts the same fundamental information across both search conditions, and this process is a key determinant of overall performance. This finding contributes to our overall understanding of color-shape conjunction search performance, which has often been described as more complex than feature search performance (e.g., Treisman & Gelade[1]). While early theories like Feature Integration Theory proposed a fundamental distinction between feature and conjunction search (i.e., one can rely on parallel detection while the other requires serial scanning), subsequent models such as Guided Search[34], Attentional Engagement Theory (AET)[35], and computational implementations like the SERR model[36], have tried to bring about a more nuanced discussion of the two types of search. The fact that the winning model in the present study was the Parallel-Sequential Rejection model adds to this literature by highlighting the importance of considering different distractor processing strategies. Specifically, while distractors in feature search are processed and rejected simultaneously, color-shape conjunction search seems to lead to a more hierarchical type of processing, with distractors evaluated and rejected by subsets, each governed by its own rejection rate (or similarity-relation to the target, as assessed by peripheral vision). This study extends prior work by providing computational evidence supporting this processing architecture and situating it within the broader context of search models.

It is important to note that a parallel-sequential mode of distractor rejection is not necessarily observed when different types of distractors are intermixed in the display. Indeed, across multiple studies, the Parallel-Simultaneous Rejection Model (where all distractors, regardless of type, are rejected simultaneously, each at their own rate) had been the best-performing predictive model across a range of stimulus types and complexity (real-life objects[12]; simple geometric shapes[13]; oriented lines[14]). Thus, distractor heterogeneity per se does not trigger the use of a parallel-sequential rejection mode. Target Contrast Signal Theory provides a framework for understanding this. It appears that the very nature of the conjunction search task might be responsible for the adoption of this novel rejection strategy: because all distractors share one of the target-defining features with the target, conjunction search displays create conditions where at every location, there is one evidence accumulation process that is failing to detect a difference between the distractor at that location and the target template (i.e., the feature that is identical to that of the target does not provide any positive evidence of being different from the target). This is a critical difference with previous studies using heterogeneous displays and/or multidimensional objects: across those studies, the target and distractor stimuli never shared any basic visual features (e.g., they differed along color and shape in Lleras et al[13]). Perhaps because evidence does not gather simultaneously across both target-defining features at all locations, participants switch to a processing mode where they first attempt to reject one subset of items based on a positive accumulation of evidence (e.g., reject all the items that have a different color than the target template), before switching to evaluating the subset of remaining items along a second featural difference (i.e., reject all the remaining items that have a different shape than the target template). This account is consistent with the findings from Rutishauser and Koch[26], who analyzed eye movements during conjunction search and found evidence for a hierarchical mode of processing, starting with attention to color, then size and orientation. Our results are also in line with other studies with both human and animals showing that, in conjunction search, observers exhibit an initial preference to attend color-defined subsets of items in the display[27–29,37].

The present results showed a consistent slowing down effect from intermixing distractors (reflected by $\beta > 1$) in both rectangular and circular grids. The slowing down effect was weakened when the distractors were segregated (Experiments 2C and 2D) compared to when the distractors were intermixed (Experiments 2A and 2B). Comparing the results of the two grids in intermixed conditions, circular grids with cortical magnification (Experiment 2B) did not significantly eliminate the conjunction search difficulty compared to rectangular grids (Experiment 2A). Furthermore, performance in color-shape conjunction search was predicted by parameters estimated in feature search, revealing there are not fundamentally different signals driving performance in feature and conjunction search tasks. The perceptual comparisons (red vs orange; circle vs triangle) remain fundamentally the same. Together, this evidence suggests that pooling-mediated processing does not fundamentally alter the comparisons that the visual system is doing to find the target, and thus, it is unlikely to be a critical limiting factor in conjunction search difficulty, at least with the stimuli and arrangements used here. This conclusion is consistent with that of Poder and Kosiło[23], who modified the traditional conjunction search design to control for crowding effects by spacing out stimuli. Using a Signal Detection analysis methodology, their results suggested that even when crowding effects were alleviated, performance in conjunction search continued to be capacity limited, suggesting crowding was not the principal factor limiting performance in these tasks (as proposed by Rosenholtz and colleagues[8,38]).

Furthermore, the Parallel-Sequential model was able to successfully account for an overwhelming amount of the variability in the data across all display conditions (all $R^2 > 0.94$), even though it did not include any parameters to model pooling-mediated processing, crowding or spatial segregation. Nonetheless, we did observe significant differences in the β parameter as a function of display and segregation conditions. These results suggest that the presence of crowding and the spatial arrangement of stimuli in the display more generally appear to dictate the degree of slowdown observed in heterogeneous search (i.e., the magnitude of the β parameter), rather than the type of information that gets extracted from the display.

Regarding the spatial arrangement of the distractors, local distractor heterogeneity was found to be a key determinant factor in slowing down performance: there was a consistently stronger slowing down effect (reflected by larger β values) when distractors were intermixed (2A and 2B) compared to when they were spatially segregated (2C and 2D). This was observed across both types of display configuration (rectangular versus circular grids). These results are consistent with the proposal in Lleras et al.[13]that the strength of inter-item interactions is a key factor impacting search efficiency[13,14,39]. Lleras et al.[13] proposed that inter-item interactions can facilitate distractor rejection: when nearby objects are of the same type (i.e., high distractor-distractor similarity), search performance is improved because all nearby accumulators are simultaneously accumulating identical evidence, so cross-talk between accumulators improves performance of all accumulators; in contrast, when nearby objects are of different types (i.e., low distractor-distractor similarity), search performance is comparatively slower, because the evidence being accumulated at any one location is different than the evidence being accumulated at nearby locations. In other words, the evidence collected at any one location does not inform what evidence might there be at nearby location, thus, not helping nearby accumulators.

A related consideration within the framework of Guided Search (GS)[40] is whether segregating distractors could increase bottom-up contrast, thereby making the contrast between the target and its surrounding distractors higher. How would such changes in bottom-up contrast between the intermixed and segregated displays impact our findings? To start, let us define bottom-up contrast as a measure of signal difference between an item and its immediate surroundings. Now, let us work through the logic. Imagine a display consisting of four items, one in each quadrant of the visual field. Clearly, at this separation, the contrast or bottom-up signal of each item is solely dependent on its own features and how much they differ from the background. In other words, at small set sizes, and given how we distributed items in the display, we can be fairly confident that the bottom-up signal of items is relatively unaffected by the segregation manipulation. As set size increases, there are higher chances that two items will be near each other in our search grid, and might impact each other's bottom-up contrast. That said, the minimal separation we implemented in the crowding-minimized displays is such that one can be fairly confident that even two nearby items will not impact each other's bottom-up contrast. At least, that is what is suggested by the pooling-mediated processing literature. Thus, segregated displays might create higher contrast signals for targets, but only under conditions with large set sizes and with small inter-item spacing. We can refer to this as a (new) factor called display density. Interestingly, our modeling contains no implementation for a "display density" parameter and the winning model still captures the overwhelming majority of the variance in the tasks (above 94% of variance). To be sure, it is entirely possible that creating a model of local density and applying it to the data might indeed further improve the predictions of the models. However, such modeling would be out of scope of the present paper. Importantly, because our data will be made public upon publication, interested researchers will be able to investigate this issue by adding a factor that tracts local density around the target in our displays and attempt further modeling.

In addition to pooling-mediated processing, the concept of functional viewing field (FVF) has also been recently proposed as a key factor in determining search performance. Hulleman and Olivers[41] proposed that the degree of target-distractor similarity effectively limited the region of space over which the visual system can make accurate perceptual discriminations during search. When target-distractor similarity is very low, then in a single glance, observers can process all the items in the display and find the target. The FVF in this case would be the entire display. At medium levels of target-distractor similarity, the FVF shrinks to encompass only a subset of items located around the current point of fixation. More than one eye movement might then be required to find the target, in these cases, basically until the target lands inside an FVF. At very high levels of target-distractor similarity, the FVF is effectively one item: only the currently fixated item is evaluated and search becomes extremely serial, as dictated by this slow progression of single-item FVFs across the scene. This concept has proven quite useful in understanding visual search, in particular, the role of eye movements in determining search times. This approach, however, is limited to target-distractor differences that reduce the FVF to some extent smaller than the whole displays. That is not the case in feature search, certainly not in the feature search conditions which we chose here, as evidenced by the logarithmic RT patterns observed in those tasks (a signature of parallel processing), and relatively small search slope values. One shortcoming of the FVF account of search is that it does not provide a framework for understanding distractor heterogeneity effects. It is not clear why the FVF in a conjunction search task is not the entire display, given that each featural difference (along color and along shape) does allow for whole-display FVFs in the corresponding feature search task.

More recently, Wolfe[40] proposed the existence of three distinct functional visual field concepts that at play during search: the resolution FVF, exploratory FVF, and attentional FVF. The spatial segregation manipulation in our experiments altered the nature of information processed within each fixation and here, the concept of attentional FVF might help explain why participants search slower in heterogeneous displays. The attentional FVF governs covert deployments of attention during fixation and encapsulates the capacity to process multiple items within a single visual intake. In scenarios where items are intermixed, the complexity within the attentional FVF increases, potentially slowing down attention movements between attended items and potentially reducing the extent of the attentional FVF. In other words, as heterogeneity slows down movements of attention inside the attentional FVF, it ends up reducing the effective size of it, as the eyes will end up moving to a new location, having processed fewer items.

Finally, returning to the early visual attention theories attributing the slowdown of conjunction search performance to the cognitive demands of feature binding, our findings offer a distinct perspective. The decrease in search efficiency observed in our studies does not appear to stem from an inherent failure to bind features, but rather from the strategic adjustments necessitated by the task's design and the nature of the stimuli. This insight provides a pivotal shift in understanding the underlying mechanisms of conjunction search difficulties. Our experiment results demonstrating a Parallel-Sequential Rejection mode of processing distractors, suggested that the real challenge lies not in the binding per se but in how the visual system manages the information presented across the visual field. Rather than a parallel, simultaneous processing of all elements typically observed in efficient and parallel visual search tasks like feature searches, a strategic, sequential attention to subsets of features is deployed to deal with the unique design of color-shape conjunction search stimuli: compound stimuli that always match the target in one of the two target-defining features. This approach, wherein participants systematically process subsets of distractors that share features with the target, reflects a more nuanced interaction with the visual display. It is a strategy prompted not by the inability of the system to bind features effectively but by the need to optimize processing in a context where features are shared across both targets and distractors.

Limitations on generalizability

First, we must acknowledge that our study focused on two specific features: color and shape. It may be premature to generalize our findings to any other two sets of features that could be used to define a conjunction search task. Indeed, some may argue that the use of color in this study might be problematic because color may have a privileged role in terms of its impact on attention and perceptual grouping. Thus, one might wonder whether the fact that the Parallel-Sequential model won here might simply reflect a natural tendency in observers to always first use color to reject some items in the display, before engaging with other available features. That said, on a number of previous published experiments where color could have been used to guide a sequential rejection of distractors, we found evidence for Parallel-Simultaneous (rather than Parallel-Sequential) rejection (see[13]). For example, in Lleras et al.[13], participants were asked to find a red triangle in a display containing blue circles (low color and shape similarity) and orange diamonds (high color and shape similarity). In theory, participants could have chosen to engage in color-based Parallel-Sequential rejection of distractors: first rejecting the blue items and then the orange diamonds. But, they did not do so, choosing instead to reject both blue circles and orange diamonds simultaneously. Furthermore, in a separate study investigating a color-color conjunction search task[31], we found evidence of Parallel-Sequential Rejection being at play. In sum, the (admittedly limited) available evidence does seem to suggest that the Parallel-Sequential nature of distractor rejection in color-shape conjunction search tasks might be caused by the nature of the conjunction search task, rather than by the specific features we chose for this study. More studies will be needed to obtain further validation for this interpretation.

As a word of caution, it is worth noting that the β parameter in Equation (4) is not solely determined by interitem interactions. As a multiplicative parameter, other factors that systematically impact slope parameters (D) will load in β, such as differences in eccentricity[19], crowding[42], and stimulus size.

We do not expect our findings to generalize to real-world scenarios, as it is difficult (or at least very rare) to find natural visual environments where the target is defined by two visual features and shares one of those features with every object in the scene. However, we do expect it to generalize to other conjunction search conditions, using different stimuli and different visual dimensions.

Data availability

The datasets generated and/or analysed during the current study are available in the Open Science Framework repository, link: https://osf.io/ymv28/.

Received: 3 October 2024; Accepted: 26 February 2025
Published online: 05 March 2025

References

1. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cognitive psychology* **12**, 97–136 (1980).
2. Eckstein, M. P. The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological science* **9**, 111–118 (1998).
3. Eckstein, M. P., Thomas, J. P., Palmer, J. & Shimozaki, S. S. A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & psychophysics* **62**, 425–451 (2000).
4. Huang, L. & Pashler, H. Attention capacity and task difficulty in visual search. *Cognition* **94**, B101–B111 (2005).
5. Buetti, S., Xu, J. & Lleras, A. Predicting how color and shape combine in the human visual system to direct attention. *Scientific reports* **9**, 20258 (2019).
6. Xu, Z. J., Lleras, A. & Buetti, S. Predicting how surface texture and shape combine in the human visual system to direct attention. *Scientific reports* **11**, 6170 (2021).
7. Balas, B., Nakano, L. & Rosenholtz, R. A summary-statistic representation in peripheral vision explains visual crowding. *Journal of vision* **9**, 13–13 (2009).
8. Rosenholtz, R., Huang, J., Raj, A., Balas, B. J. & Ilie, L. A summary statistic representation in peripheral vision explains visual search. *Journal of vision* **12**, 14–14 (2012).
9. Portilla, J. & Simoncelli, E. P. A parametric texture model based on joint statistics of complex wavelet coefficients. *International journal of computer vision* **40**, 49–70 (2000).
10. Townsend, J. T. & Ashby, F. G. *Stochastic modeling of elementary psychological processes* (Cambridge University Press, New York, 1983).
11. Buetti, S., Cronin, D. A., Madison, A. M., Wang, Z. & Lleras, A. Towards a better understanding of parallel visual processing in human vision: Evidence for exhaustive analysis of visual information. *Journal of Experimental Psychology: General* **145**, 672 (2016).
12. Wang, Z., Buetti, S. & Lleras, A. Predicting search performance in heterogeneous visual search scenes with real-world objects. *Collabra: Psychology* **3**, 6 (2017).
13. Lleras, A., Wang, Z., Madison, A. & Buetti, S. Predicting search performance in heterogeneous scenes: Quantifying the impact of homogeneity effects in efficient search. *Collabra: Psychology* **5**, 2 (2019).
14. Xu, Z. J., Lleras, A., Shao, Y. & Buetti, S. Distractor-distractor interactions in visual search for oriented targets explain the increased difficulty observed in nonlinearly separable conditions. *Journal of Experimental Psychology: Human Perception and Performance* **47**, 1274 (2021).
15. Lleras, A. et al. A target contrast signal theory of parallel processing in goal-directed search. *Attention, Perception, & Psychophysics* **82**, 394–425 (2020).
16. Duncan, J. & Humphreys, G. W. Visual search and stimulus similarity. *Psychological review* **96**, 433 (1989).
17. Xu, Z. J., Ng, G., Lleras, A., Hummel, J. E. & Buetti, S. How does multidimensional complexity impact processing efficiency in visual search?. *Journal of Vision* **21**, 2168–2168 (2021).
18. Pelli, D. G. Crowding: A cortical constraint on object recognition. *Current opinion in neurobiology* **18**, 445–451 (2008).
19. Wang, Z., Lleras, A. & Buetti, S. Parallel, exhaustive processing underlies logarithmic search functions: Visual search with cortical magnification. *Psychonomic bulletin & review* **25**, 1343–1350 (2018).
20. Rovamo, J. & Virsu, V. An estimation and application of the human cortical magnification factor. *Experimental brain research* **37**, 495–510 (1979).
21. Dugué, L., Xue, A. M. & Carrasco, M. Distinct perceptual rhythms for feature and conjunction searches. *Journal of Vision* **17**, 22–22 (2017).
22. McElree, B. & Carrasco, M. The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance* **25**, 1517 (1999).
23. Põder, E. & Kosiło, M. What limits search for conjunctions of simple visual features?. *Journal of Vision* **19**, 4–4 (2019).
24. Treisman, A. & Sato, S. Conjunction search revisited. *Journal of experimental psychology: human perception and performance* **16**, 459 (1990).
25. Zhuang, X. & Papathomas, T. V. Cue relevance effects in conjunctive visual search: Cueing for location, color, and orientation. *Journal of Vision* **11**, 6–6 (2011).
26. Rutishauser, U. & Koch, C. Probabilistic modeling of eye movement data during conjunction search via feature-based attention. *Journal of Vision* **7**, 5–5 (2007).
27. Hannus, A., Van den Berg, R., Bekkering, H., Roerdink, J. B. & Cornelissen, F. W. Visual search near threshold: Some features are more equal than others. *Journal of Vision* **6**, 15–15 (2006).
28. Egeth, H. E., Virzi, R. A. & Garbart, H. Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance* **10**, 32 (1984).
29. Kaptein, N. A., Theeuwes, J. & Van der Heijden, A. Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance* **21**, 1053 (1995).
30. Wang, Y., Buetti, S. & Lleras, A. Search for color x color conjunction in heterogeneous display can unfold in parallel. *Journal of Vision* **22**, 4087–4087 (2022).
31. Wang, Y., Buetti, S., Cui, A. Y. & Lleras, A. Color-color feature guidance in visual search. *Attention, Perception, & Psychophysics* (in press).
32. Zelinsky, G. J. A theory of eye movements during target acquisition. *Psychological review* **115**, 787 (2008).

33. Wolfe, J. M. Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review* **1**, 202–238 (1994).
34. Wolfe, J. M., Cave, K. R. & Franzel, S. L. Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human perception and performance* **15**, 419 (1989).
35. Duncan, J. & Humphreys, G. Beyond the search surface: visual search and attentional engagement. *Journal of Experimental Psychology: Human perception and performance* (1992).
36. Humphreys, G. W. & Muller, H. J. Search via recursive rejection (serr): A connectionist model of visual search. *Cognitive Psychology* **25**, 43–110 (1993).
37. Motter, B. C. & Belky, E. J. The guidance of eye movements during active visual search. *Vision research* **38**, 1805–1815 (1998).
38. Rosenholtz, R., Huang, J. & Ehinger, K. A. Rethinking the role of top-down attention in vision: Effects attributable to a lossy representation in peripheral vision. *Frontiers in psychology* **3**, 13 (2012).
39. Cui, A. Y., Lleras, A., Ng, G. J. P. & Buetti, S. Complex background information slows down parallel search efficiency by reducing the strength of interitem interactions. *Journal of Experimental Psychology: Human Perception and Performance* **49**, 1053 (2023).
40. Wolfe, J. M. Guided search 6.0: An updated model of visual search. *Psychon. Bull. Rev.* **28**, 1060–1092 (2021).
41. Hulleman, J. & Olivers, C. N. The impending demise of the item in visual search. *Behavioral and Brain Sciences* **40**, e132 (2017).
42. Madison, A., Lleras, A. & Buetti, S. The role of crowding in parallel search: Peripheral pooling is not responsible for logarithmic efficiency in parallel search. *Attention, Perception, & Psychophysics* **80**, 352–373 (2018).

Author contributions

S.B., A.Y.C., and A.L. contributed to conceptualization. S.B., A.Y.C., and A.L. contributed to the design and methodology. A.Y.C. contributed to software, investigation, data analysis, and visualization. Z.J.X. contributed to data analysis. S.B., A.Y.C., and A.L. contributed to writing. This project was supported by an NSF grant to S.B. (Award Number BCS 1921735).

Funding

This project was supported by an NSF grant to S.B. (Award Number BCS 1921735).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92453-3.

Correspondence and requests for materials should be addressed to A.Y.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.