Cognition 182 (2019) 8–13

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Gestalt similarity groupings are not constructed in parallel

Dian Yu*,[1], Derek Tam, Steven L. Franconeri

Department of Psychology, Northwestern University, United States

ARTICLE INFO

Keywords:
Grouping
Perceptual organization
Visual attention
Feature-based selection
Visual search

ABSTRACT

Our visual system organizes spatially distinct areas with similar features into perceptual groups. To better understand the underlying mechanism of grouping, one route is to study its capacity and temporal progression. Intuitively, that capacity seems unlimited, and the temporal progression feels immediate. In contrast, here we show that in a visual search task that requires similarity grouping, search performance is consistent with serial processing of those groups. This was true across several experiments, for seeking a single ungrouped pair among grouped pairs, vice versa, and for displays with tiny spacings between the grouped items. In a control condition that ruled out display complexity confounds, when the small inter-object spacing was removed so that that pairs touched, removing the need to group by similarity, search became parallel. Why is similarity grouping so slow to develop? We argue that similarity grouping is 'just' feature selection - seeing a red, bright, or square group *is* global selection of those features. This account predicts serial processing of one feature group at a time, and makes new falsifiable predictions about how properties of feature-based selection should be reflected in similarity grouping.

1. Introduction

Our visual system groups visually separated areas of the visual field that share certain features, such as color, shape, or orientation. Grouping by similarity is one of the Gestalt grouping rules that describe how we extract structure from the visual world. Despite more than a century of perceptual grouping research (Ehrenfels, 1890), we have little understanding of the mechanism that underlies similarity grouping, in contrast to other grouping mechanisms, such as grouping by collinearity, where plausible neural mechanisms have been proposed (Roelfsema, 2006).

When seeking mechanisms that underlie a perceptual processes, we often seek clues by studying its properties: what kinds of stimuli engage it, whether the process has a capacity limitation, and its temporal progression. For this last property, one major division is between ‘fast’ and ‘slow’ processes (Holcombe, 2009), such as quickly recognizing the existence of two colors, versus judging their relative spatial relations (Logan, 1994, 1995), or binding local elements into global form versus binding global form with color (Clifford et al., 2004). These slower processes are often assumed to require some degree of attention – serial processing of selected aspects or parts of the stimulus (Franconeri, 2013; Holcombe, 2009; but see VanRullen, 2009 for alternatives).

Intuitively, similarity grouping feels fast and spontaneous. According to one of the founders of Gestalt psychology, “I stand at the window and see a house, trees, sky. Theoretically I might say there were 327 brightnesses and nuances of colour. Do I *have* 327? No. I have sky, house, and trees…The concrete division which I *see* is not determined by some arbitrary mode of organization lying solely within my own pleasure; instead I see the arrangement and division which is given there before me.” (Wertheimer, 1923).

Although similarity grouping feels fast, some evidence suggests that more time is needed to construct more groups. For example, similarity groups do not seem to automatically arise for multiple feature dimensions at once. One study showed that when similarity cues compete – with shape suggesting one grouping structure and color another – viewers tend to use one cue or the other in a bimodal fashion, instead of a combination of the two (Huang, 2015). However, these results are based on perceptual reports, not performance metrics. Other evidence suggestions that as a task requires the viewer to construct more color-similarity groups, response times increase. For example, symmetry judgments of a multicolored pattern were slower when the pattern was more heterogeneously colored (Huang & Pashler, 2002; Wagemans, 1995), suggesting that it took more time to process greater number of color groups. But one might instead argue that the slowdown was due to the increased complexity of the display. Another study showed that viewers found it difficult to locate common motion groups (which are

---

* Corresponding author at: Northwestern University, Department of Psychology, 2029 Sheridan Rd, Evanston, IL 60208, United States.
E-mail address: dianyu2017@u.northwestern.edu (D. Yu).

[1] Present address: Room D542, 32 Vassar St, Cambridge, MA 02139, United States.

https://doi.org/10.1016/j.cognition.2018.08.006
Received 2 October 2016; Received in revised form 9 August 2018; Accepted 9 August 2018

arguably a form of similarity grouping if motion direction is treated as a feature) among non-common-motion groups in a visual search task (Levinthal & Franconeri, 2011). Nonetheless, this study could be open to the critique that common-motion grouping is a special case – indeed, it is typically categorized as an independent type of grouping that includes common motion direction and common luminance changes (Brooks, 2013; Sekuler & Bennett, 2001).

In the current study, we provide the strongest evidence to date that similarity grouping is a slow process, by showing that color, typically the strongest similarity grouping cue, does not support grouping in parallel. In a series of visual search tasks, we repeatedly find inefficient visual search when seeking groups among non-groups (and vice-versa) when similarity grouping is required. These costs disappear when targets can be identified in a process that does not involve similarity grouping: by shifting objects by a few pixels so that their edges touch, viewers can use color contrast edges as more efficient proxies to locate the target, demonstrating that the inefficient search in the grouping condition cannot be due to confounded factors of higher eccentricity or display complexity at higher set sizes. In the General Discussion, we argue for a concrete and parsimonious mechanism that explains why similarity grouping is slow: that similarity grouping is 'just' serial feature selection. Seeing a red, bright, or square group is global selection of those features, and this process can only occur for one feature value at a time.

2. Experiment 1

Experiment 1 shows that search for a single ungrouped color pair among color pairs is serial when similarity grouping is required, but it is parallel when objects touch.

2.1. Method

2.1.1. Participants

Fifteen Northwestern University undergraduates and members of the Evanston, IL community signed an informed consent to participate in this study in exchange for monetary compensation. All participants were required to have normal or corrected-to-normal acuity and normal color vision. The human subjects review board at the Northwestern University approved all procedures.

2.1.2. Apparatus

All experiments were controlled via MATLAB with the Psychophysics toolbox (Brainard, 1997; Kleiner et al., 2007) on an Apple Mac Mini running the OS X operating system. The displays were 16-inch CRT monitors with a resolution of 1024 × 768 pixels. Participant head movement was unrestrained, but the average viewing distance was approximately 50 cm, so that screen area subtended 36.0 × 27.3 degrees of visual angle (37.5px/degree).

2.1.3. Stimuli and procedure

We generated a set of roughly perceptually equiluminant colors (Eisner & MacLeod, 1981) against a gray background (RGB code #808080) for use in this experiment. Pairs of squares were arranged into two concentric circles surrounding the center fixation (see first panel in Fig. 1 for illustration). Each display contained between 3 and 7 pairs of squares (randomized across trials) among which only one pair consisted of two different colored squares. All the pairs were either touching (in the touching condition) or separated by a small gap (in the separated condition). These two types of trials randomly interleaved within the same block. Participants were instructed to search for the broken pair in both conditions.

Specifically, each square was 0.267 degrees in length and width. The centers of the outer squares were 2.00 degrees from the center of the fixation in both touching and separated conditions. The centers of inner squares were 1.73 degrees from the center of the screen when touching the other squares, and 1.33 degrees away when separate. The set of perceptually equiluminant colors consisted of red (RGB value: 190,0,0), green (RGB value: 49,126,0), blue (RGB value: 0,97,168), orange (RGB value: 165,84, 0), purple (RGB value: 130,0,169), brown (RGB value: 124,111,0), and cobalt (RGB value: 0, 48,145).

In a typical trial, a variable inter-trial interval (500–1500 ms) preceded each trial. Participants were instructed to press a button immediately upon locating the target group. The time between the stimulus appearing on screen and when the participant made the initial response was recorded as the response time. Following the initial button-press, a mask appeared on the display for 500 ms, followed by an answer-input screen mimicking the layout of the first screen, except with uncolored (black) squares. Participants then made an unspeeded response to indicate which pair of these squares was the ungrouped target pair. Participants completed 14 practice trials and 280 experimental trials with a break occurring every 70 trials.

2.2. Results

Participants were highly accurate in the touching (98.4%) and separated (97.3%) conditions. We analyzed correct trials with a response made within 2000 ms (96.5% of all trials). Search slopes were relatively high ($M = 30.7$ ms/pair) and greater than zero, $t(14) = 4.96$, $p < 0.001$, Cohen's $d = 1.28$, and also greater than slopes for the touching condition ($M = 3.1$ ms/pair), $t(14) = 3.82$, $p = 0.002$, Cohen's $d = 0.99$, which were not different from zero, $t(14) = 1.02$, $p > 0.250$, Cohen's $d = 0.26$.

3. Experiment 2

Experiment 1 is open to critique that the separated condition has, in some sense, twice the number of "objects" as the touching condition. Experiment 2 addresses this concern by equating the number of perceived objects.

3.1. Method

3.1.1. Participants

Eleven Northwestern University undergraduates and members of the Evanston, IL community signed an informed consent to participate in this study in exchange for monetary compensation.

3.1.2. Stimuli and procedure

The stimuli in this experiment were similar to those of Experiment 1, with the addition of a black ring 1.07 degrees thick in the display, either in front of, or beneath, the rectangles of various colors (1.39 by 0.53 degrees). The target rectangle was composed of two colors meeting at the center of the rectangle, and the distractor rectangles were uniformly colored. The number of rectangles on the screen varied from 3 to 6. The task was otherwise identical to that of Experiment 1.

3.2. Results

Accuracy in the ring-below condition (97.6%) was similar to the ring-above condition (94.9%). One participant's data were withheld from all analysis because of a low accuracy rate (62.2%). Results from the correct trials with a response made within 2000 ms were analyzed (93% of all trials). Search slopes in the ring-above condition were high ($M = 40.9$ ms/pair) and greater than zero ($t(9) = 6.32$, $p < 0.001$, Cohen's $d = 2.13$), and also greater than slopes for the ring-below condition ($M = 8.4$ ms/pair), ($t(9) = 7.49$, $p < 0.001$, Cohen's $d = 2.51$), which were not different from zero ($t(9) = 1.42$, $p = 0.188$, Cohen's $d = 0.47$)

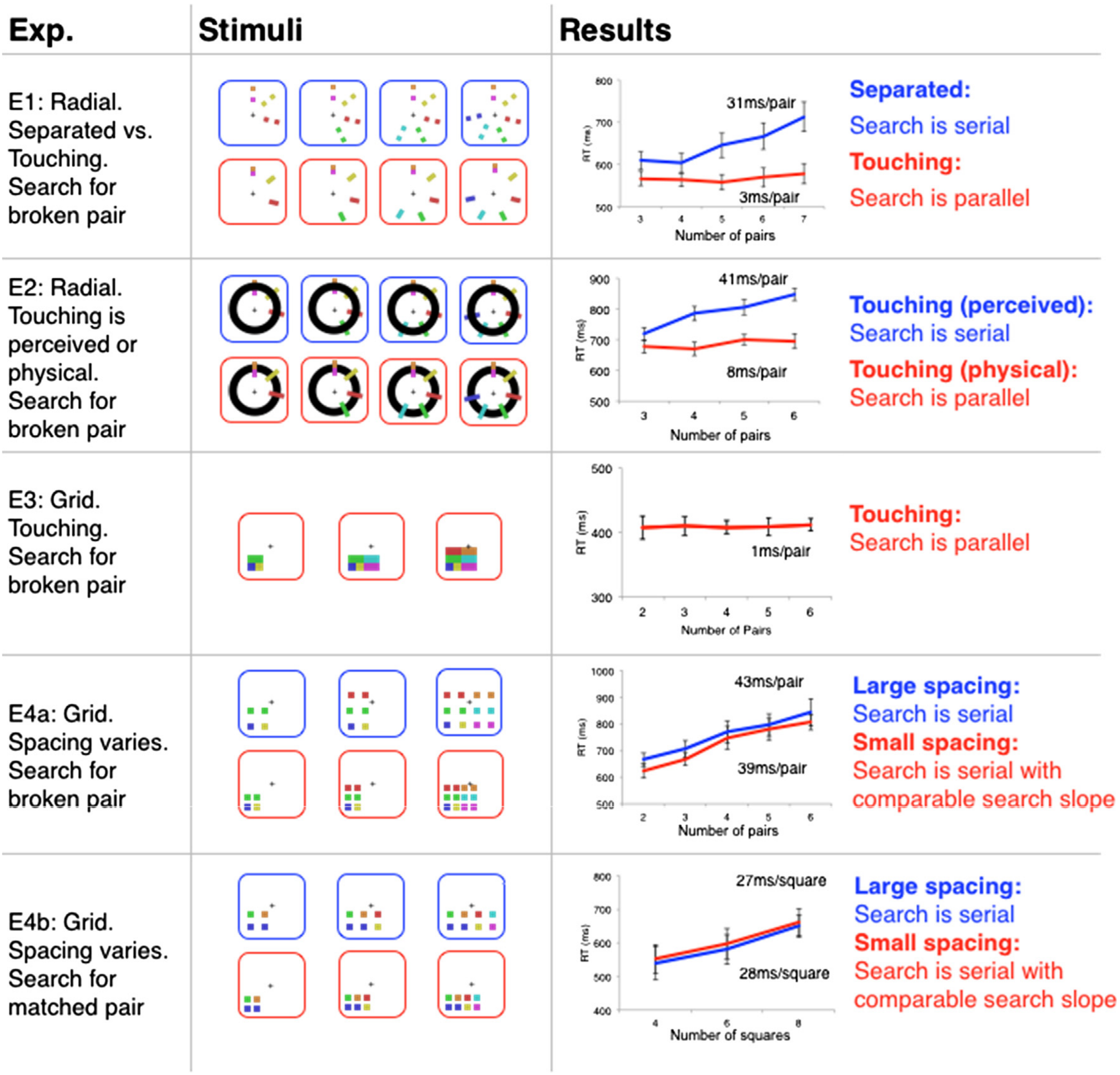

Fig. 1. Experimental displays and results from Experiment 1–4.

4. Experiment 3 & 4

The circular arrangement of Experiment 1–2 was intended to equate the eccentricity of all groups. But it is possible that this design choice made search for broken groups inefficient because all pairs were *already* moderately grouped by proximity. Experiments 3&4 therefore replicated the results with evenly spaced grid displays that are immune to this critique.

4.1. Method

4.1.1. Participants

Twelve (in Experiment 3) and thirteen (in Experiment 4) Northwestern University undergraduates and members of the Evanston, IL community signed an informed consent to participate in this study in exchange for monetary compensation.

4.1.2. Stimuli

Stimuli were pairs of 4–12 colored squares (1.1 degrees in width; identical colors as Experiment 1), arranged in a grid, creating 2–6 paired groups. The pairs of squares could be paired either horizontally or vertically, randomly selected on each trial (see Fig. 1, panel 3&4 illustrating the horizontal pairing). Squares were touching in Experiment 3, and were separated by either 1.10 degree (large spacing condition) or 0.14 degree (small spacing condition) in Experiment 4. In Experiment 4 square locations were jittered (0–12.5% of the inter-square distance).

Average object eccentricity was controlled across set sizes in both experiments. But for displays with smaller spacing between squares, including the touching condition where the spacing was absent, spacing was reduced or eliminated by pushing objects toward the *outer* edge of the screen. These more eccentric positions should make these conditions more difficult, which makes it more surprising that the touching condition shows by far the lowest RTs and search slopes.

4.1.3. Procedure

A variable inter-trial interval (1000–2000 ms) preceded each trial, containing a fixation training display: a fixation cross over a square (1.125 degrees) random-dot pattern (50% black vs. white pixels); each pixel subtending 0.035 degrees. The pattern flickered at 37.5 Hz in counter-phase reversal of black and white pixels. Under stable fixation, the pattern temporally sums to gray, while even small saccades lead to a percept of a jagged random-dot pattern. Participants were informed that the square should appear grey if they fixated on the cross, but a random-dot pattern would pop out if they move their eyes. Practicing this procedure has been shown to improve the accuracy of fixation in a subsequent experiment (Guzman-Martinez, Leung, & Franconeri, 2009).

The grid of colored squares then appeared, and participants pressed the space bar upon locating the target. After the initial button press, a mask appeared immediately for 500 ms, followed by an unspeeded response input screen with rectangular frames showing previous locations of the pairs, where they used the mouse to click on the frame at the location of the target pair.

In Experiment 3 and 4a, there was a single 'ungrouped' pair consisting of two squares of different colors, while the other pairs consisted of two squares of the same color. Participants were asked to search for the ungrouped pair. Since there was no gap between squares in Experiment 3, the display can also be seen as consisting of two adjacent squares among rectangles, so we directly asked participants to look for the adjacent squares. In Experiment 4b, participants searched for a single matched pair with adjacent squares of the same color, while the remaining squares were different colors chosen without replacement.

Experiment 3 consisted of 3 blocks with 96 trials in each block. Both Experiment 4a and 4b consisted of 3 blocks of large-spacing condition and 3 blocks of small spacing condition, 96 trials in each block. Blocks of different conditions were interleaved with their sequence counterbalanced between participants. The first blocks of all conditions in both experiments served as practice and were not included in our analysis.

4.2. Results

Accuracies in the small spacing conditions (Experiment 4a: 96.4%, Experiment 4b: 98.2%) were similar to the large spacing conditions (Experiment 4a: 96.4%, Experiment 4b: 96.6%), which are also similar to the touching condition in Experiment 3 (98.7%). Results from the correct trials with a response made within 1500 ms were analyzed (Experiment 3: 98.5%; Experiment 4a: 95.5%; Experiment 4b: 96.8% of all trials).

Search was efficient when squares touched in Experiment 3, where search slopes ($M = 0.7$ ms/pair) were not significantly different from zero ($t(11) = 0.3$, $p > 0.25$, Cohen's $d = .29$). However, when squares were separated in Experiment 4a, search became inefficient. Search slopes for both the large (42.9 ms/pair) and small spacing (39.1 ms/pair) are greater than zero (small spacing: $t(12) = 9.71$, $p < .001$, Cohen's $d = 2.70$; big spacing: $t(12) = 6.62$, $p < .001$, Cohen's $d = 1.84$). Furthermore, search slopes did not different between two spacing conditions ($t(12) = .71$, $p > .250$, Cohen's $d = .20$).

Similar results were found in Experiment 4b, where participants searched for the matched pair in a group of differ colored squares. Here we measure search slope as the response time change per new square, instead of per new pair of squares, because each new square adds a new color to the display. Search slopes for both the large (27.1 ms/square) and small spacing (27.7 ms/square) are greater than zero (small spacing: $t(12) = 6.21$, $p < .001$, Cohen's $d = 1.73$; big spacing: $t(12) = 6.62$, $p < .001$, Cohen's $d = 2.11$). Furthermore, search slopes did not differ between two spacing conditions ($t(12) = .19$, $p > .250$, Cohen's $d = .05$).

5. General discussion

Our results show that color grouping, as a case study of Gestalt similarity grouping, is a slow visual process that cannot support parallel visual search across groups. This was true even when objects appeared to group under an occluding surface, and when objects were tightly spaced. To ensure that these additional response time costs at higher set sizes were not due to confounded increases in display eccentricity or complexity, control conditions shifted objects by a few pixels so that they touched, allowing viewers to rely on 'fast' detection of color contrast edges as proxies to locate the target. Search became parallel, even though these conditions maintained the increases in display eccentricity and complexity, showing that similarity grouping, per se, that is slow.

Complex visual processing can unfold instantaneously in some cases, such as scene recognition (125 ms, Potter, 1975) or gender discrimination of faces (can be completed when attention was distracted by a dual task, Reddy, Wilken, & Koch, 2004). But the present results are consistent with evidence that some types of flexible extraction of visual structure can be a surprisingly slow (Franconeri, 2013; Holcombe & Cavanagh, 2001). Extracting a spatial relationship (e.g. gray to the left of black) is strikingly inefficient in various search tasks (Huang & Pashler, 2005; Logan, 1994, 1995; Palmer, 1994; Reddy & VanRullen, 2007; Wolfe, 1998). Some accounts of this slowdown point to the need to select visual information serially over time. For spatial relationships, sequential processing of each object (black; grey) may be needed to extract each object's relative position to the other (Franconeri, Scimeca, Roth, Helseth, & Kahn, 2012).

5.1. A potential mechanism for similarity grouping

Why might similarity grouping unfold slowly over time? One recent proposal for the mechanism of similarity grouping predicts that creating similarity groups should be in fact slow, and actually strictly serial (Huang & Pashler, 2007; Levinthal & Franconeri, 2011). This proposal argues similarity groups are constructed by feature-based selection (Huang & Pashler, 2007; Levinthal & Franconeri, 2011). Feature-based attention is known to weigh information from locations in the visual field that contain a given color, shape, orientation, or motion direction (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Liu, Larsson, & Carrasco, 2007; Saenz, Buracas, & Boynton, 2002). Co-activation of these disparate locations could produce the feeling that they 'belong' together, and allow global operations over those locations (Huang & Pashler, 2007; Saenz, Buraĉas, & Boynton, 2003).

One reason to doubt this account is that it makes a counterintuitive prediction: that similarity groups *must* be constructed in a strictly serial fashion. If red objects are grouped because a viewer is selecting for red regions of the visual field, then green regions must not be selected, or they would be grouped with the red objects (Huang & Pashler, 2007). Such one-at-a-time grouping might seem incongruent with our perception of a detailed visual representation, including multiple instances of similarity groups. The present data are entirely consistent with this prediction.

At first glance, the idea that feature selection is needed to create similarity groups might appear to contradict evidence that similarity groups can be constructed 'without attention' in a dual-task paradigm that focused spatial attention elsewhere (Kimchi & Razpurker-Apfeld, 2004; Moore & Egeth, 1997; Russell & Driver, 2005; Shomstein, Lee, & Behrmann, 2010). But these effects rely on different definitions of 'attention' – if a participant completes a dual task that requires spatial selection of one region of a display, it would not rule out their ability to tune feature-based attention toward other parts of a display.

5.2. Why then, does similarity grouping feel fast and parallel?

Wertheimer (1923) described a structured world where similarity groups seemed to exist in parallel – how could a serial mechanism of similarity grouping support such a rich percept? The feeling of simultaneous perception of similarity groups may be an illusion that is subserved by rapid and on-demand construction of visual information (e.g., Rensink, 2000). It is not unlike the illusion that the light of the fridge is always on – every time you check for the presence of a certain similarity group, it is immediately constructed on demand, giving you the illusion that you could see it all the time.

If similarity grouping is serial, then how does the visual system know which feature to group, and in what order? One route could be that when a single object is attended, its constituent features (e.g., red, circle, vertical stripes) are selected as well (Katzner, Busse, & Treue, 2009; Melcher, Papathomas, & Vidnyánszky, 2005; Schoenfeld, Hopf, Merkel, Heinze, & Hillyard, 2014). This co-selection of features is not confined to a single attended location. Instead, enhancement of selected features may extend across the visual field to objects with the same features (Lustig & Beck, 2012). Selecting a red object could cause a viewer to select all other red objects as well, which would cause them to group. Another route might be that our initial snapshot of a display or scene is unselective – and ungrouped by similarity – but this unselective stage produces a histogram of the features available for selection (Haberman & Whitney, 2012; Halberda, Sires, & Feigenson, 2006). Such global feature summary information, such as a forest consisting of mostly green and brown color and vertical orientations, can be acquired within a presentation as short as 30 ms (Oliva & Torralba, 2006). Knowing the available features, and their quantity, is not the same as representing groups – that requires knowing where these features are, how they are arranged, and that they are associated. This histogram may offer guidance for that grouping process. One feature might initially be chosen from this histogram (perhaps the most or least prevalent, or otherwise salient), followed by rapid shifts through feature space, according to an emergent competitive process taking input from the statistics of the display, task constraints, and previous experience.

Despite Wertheimer's claim of a single possible organization, you may feel that you can see different kinds of similarity groups in a single display, for example, a visual scene may be organized differently if one groups visual areas based on similar color versus similar texture. One might need to group a visual display according to different cues for different task. Such flexibility could be enabled by feature-based attention.

5.3. Other approaches to test the feature-based selection account

The counterintuitive idea that feature-based attention underlies similarity grouping should be supported by further converging evidence, because no single paradigm or result will be sufficient. For example, inefficient search can be due to difficulties in deciding among potential targets, instead of difficulties in constructing them (Palmer, 1995). Our 'touching' control condition minimizes this worry for the present experiments. But future work could marshal further evidence for slow similarity grouping being tied strictly to the grouping process *per se,* by testing with 'simultaneous/sequential' paradigm (Attarha, Moore, & Vecera, 2014). In addition, we recently found evidence for one-at-a-time similarity grouping using a different paradigm, by showing that color groupings (and other similarity cues) do not bias viewers' estimate of their quantity, while grouping objects by proximity, connectedness or common region does bias quantity estimates (Yu & Franconeri, in press).

In summary, we present robust evidence that color grouping shares the same characteristics and limitation of feature-based selection. Our current finding, combined with previous research, supports the parsimonious account that grouping by similarity *is* feature-based selection. This account allows us to make further predictions about the characteristics of similarity grouping with our existing knowledge of feature-based selection. For example, selection is more difficult when it is between more psychologically similar sets (Duncan & Humphreys, 1989). Does the magnitude of such similarity effects for a selection task (e.g., what are the respective difficulties of selecting between the pairwise combinations of feature value A, B, C, etc.?) correlate with the similarity effects on grouping (e.g., how strongly does A group in the presence of B, C, …?). Does feature redundancy (Nothelfer, Gleicher, & Franconeri, 2017) (e.g., select the green squares among red circles) improve feature selection in the same ways across feature selection and similarity grouping tasks. If feature-based selection elicits surround suppression in feature space (Störmer & Alverez, 2014), could the same effects be found in similarity grouping? The feature-based attention account of Gestalt similarity grouping is not only parsimonious – using a known selection mechanism to explain grouping – but it also makes falsifiable predictions for functional similarities in the way that the two processes operate.

Supplementary materials

All data analyzed in the current study are archived on Harvard Dataverse, and can be accessed at: Yu, Dian; Derek Tam; Steven Franconeri, 2018, "Data for Gestalt similarity groupings are not constructed in parallel", https://doi.org/10.7910/DVN/OD56OT, Harvard Dataverse, V1, UNF:6:utbuZXhj1ez146nqxIt0yQ = =.

Author contributions

All authors contributed to developing the study concept and the study design. Data collection and analysis were performed by D. Tam and D. Yu. D. Yu drafted the manuscript and S. Franconeri provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Cathleen Moore, John Palmer, Satoru Suzuki, and Marcia Grabowecky for helpful discussion, and Jennifer Corbett, Justin Halberda, and an anonymous reviewer for valuable comments, and William Cook for assistance in data collection. This work was supported by IIS-1162067 and BCS-1056730.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cognition.2018.08.006.

References

Attarha, M., Moore, C. M., & Vecera, S. P. (2014). Summary statistics of size: Fixed processing capacity for multiple ensembles but unlimited processing capacity for single ensembles. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1440.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Brooks, J. L. (2013). *Traditional and new principles of perceptual grouping: Oxford handbook of perceptual organization.* Oxford, England: Oxford University Press.

Clifford, C. W. G., et al. (2004). Rapid global form binding with loss of associated colors. *Journal of Vision, 4*, 1090–1101.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science, 248*(4962), 1556.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433.

Ehrenfels, C.v. (1890). Über gestaltqualitäten. *Vierteljahresschrift für wissenschaftliche Philosophie, 14*(3), 249–292.

Eisner, A., & MacLeod, D. I. (1981). Flicker photometric study of chromatic adaptation: Selective suppression of cone inputs by colored backgrounds. *JOSA, 71*(6), 705–718.

Franconeri, S. L. (2013). The nature and status of visual resources. *Oxford Handbook of Cognitive Psychology, 8481*, 147–162.

Franconeri, S. L., Scimeca, J. M., Roth, J. C., Helseth, S. A., & Kahn, L. E. (2012). Flexible visual processing of spatial relationships. *Cognition, 122*(2), 210–227.

Guzman-Martinez, E., Leung, P., Franconeri, S., Grabowecky, M., & Suzuki, S. (2009). Rapid eye-fixation training without eyetracking. *Psychonomic Bulletin & Review, 16*(3), 491–496 (Chicago).

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. *From Perception to Consciousness: Searching with Anne Treisman,* 339–349.

Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science, 17*(7), 572–576.

Holcombe, A. O. (2009). Seeing slow and seeing fast: Two limits on perception. *Trends in Cognitive Sciences, 13*(5), 216–221.

Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience, 4*(2), 127.

Huang, L. (2015). Grouping by similarity is mediated by feature selection: Evidence from the failure of cue combination. *Psychonomic Bulletin & Review, 22*(5), 1364–1369.

Huang, L., & Pashler, H. (2002). Symmetry detection and visual attention: A "binary-map" hypothesis. *Vision Research, 42*(11), 1421–1430.

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition, 94*, B101–B111.

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review, 114*(3), 599.

Katzner, S., Busse, L., & Treue, S. (2009). Attention to the color of a moving stimulus modulates motion-signal processing in macaque area MT: Evidence for a unified attentional system. *Frontiers in Systems Neuroscience, 3*, 12.

Kimchi, R., & Razpurker-Apfeld, I. (2004). Perceptual grouping and attention: Not all groupings are equal. *Psychonomic Bulletin & Review, 11*(4), 687–696.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36*(14), 1.

Levinthal, B. R., & Franconeri, S. L. (2011). Common-fate grouping as feature selection. *Psychological Science* 0956797611418346.

Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron, 55*(2), 313–323.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1015–1036.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28*(2), 103–174.

Lustig, A. G., & Beck, D. M. (2012). Task-relevant and task-irrelevant dimensions are modulated independently at a task-irrelevant location. *Journal of Cognitive Neuroscience, 24*(9), 1884–1895.

Melcher, D., Papathomas, T. V., & Vidnyánszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron, 46*(5), 723–729.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 339.

Nothelfer, C., Gleicher, M., & Franconeri, S. (2017). Redundant encoding strengthens segmentation and grouping in visual displays of data. *Journal of Experimental Psychology: Human Perception and Performance, 43*(9), 1667.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research, 155*, 23–36.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research, 13*, 1703–1721.

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science, 4*(4), 118–123.

Potter, M. C. (1975). Meaning in visual search. *Science, 187*(4180), 965–966.

Reddy, L., & VanRullen, R. (2007). Spacing affects some but not all visual searches: Implications for theories of attention and crowding. *Journal of Vision, 7*(2), 1–17.

Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of Vision, 4*(2), 4–14.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research, 40*(10), 1469–1487.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29*, 203–227.

Russell, C., & Driver, J. (2005). New indirect measures of "inattentive" visual grouping in a change-detection task. *Perception & Psychophysics, 67*(4), 606–623.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5*(7), 631–632.

Saenz, M., Buraĉas, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43*(6), 629–637.

Schoenfeld, M. A., Hopf, J.-M., Merkel, C., Heinze, H.-J., & Hillyard, S. A. (2014). Object-based attention involves the sequential activation of feature-specific cortical modules. *Nature Neuroscience, 17*(4), 619–624.

Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychological Science, 12*(6), 437–444.

Shomstein, S., Lee, J., & Behrmann, M. (2010). Top-down and bottom-up attentional guidance: Investigating the role of the dorsal and ventral parietal cortices. *Experimental Brain Research, 206*(2), 197–208.

Störmer, V. S., & Alvarez, G. A. (2014). Feature-based attention elicits surround suppression in feature space. *Current Biology, 24*(17), 1985–1988.

Vanrullen, R. (2009). Binding hardwired versus on-demand feature conjunctions. *Visual Cognition, 17*(1–2), 103–119.

Wagemans, J. (1995). Detection of visual symmetries. *Spatial Vision, 9*(1), 9–32.

Wertheimer, M. (1923). Laws of organization in perceptual forms. *A Source Book of Gestalt Psychology*.

Wolfe, J. M. (1998). What can 1,000,000 trials tell us about visual search? *Psychological Science, 9*(1), 33–38.

Yu, D., & Franconeri, S. L. (2018). Similarity grouping as feature-based selection. *Psychological Science* (in press).