CrossMark

Measuring the time course of selection during visual search

Evan M. Palmer [1,2,3] · Michael J. Van Wert [1] · Todd S. Horowitz [1,2] · Jeremy M. Wolfe [1,2]

Abstract
In visual search tasks, observers can guide their attention towards items in the visual field that share features with the target item. In this series of studies, we examined the time course of guidance toward a subset of items that have the same color as the target item. Landolt Cs were placed on 16 colored disks. Fifteen distractor Cs had gaps facing up or down while one target C had a gap facing left or right. Observers searched for the target C and reported which side contained the gap as quickly as possible. In the absence of other information, observers must search at random through the Cs. However, during the trial, the disks changed colors. Twelve disks were now of one color and four disks were of another color. Observers knew that the target C would always be in the smaller color set. The experimental question was how quickly observers could guide their attention to the smaller color set. Results indicate that observers could not make instantaneous use of color information to guide the search, even when they knew which two colors would be appearing on every trial. In each study, it took participants 200–300 ms to fully utilize the color information once presented. Control studies replicated the finding with more saturated colors and with colored C stimuli (rather than Cs on colored disks). We conclude that segregation of a display by color for the purposes of guidance takes 200–300 ms to fully develop.

Keywords Visual search · Attention: Selective · Attention

Introduction

When we search for items in the visual world, we typically use known features of the target item to limit that search. Thus, in a search for a banana, we would guide our attention to yellow, curved items of banana size. There would be no point in attending to the toaster, as it lacks the features of a banana. Bringing this into the lab, Egeth, Virzi, and Garbart (1984) measured the reaction time (RT) required to find a specific letter in an array of other letters. If observers were told that the target letter was red, they were able to search much more efficiently, apparently guiding attention only to those items that had the target color (see also Kaptein, Theeuwes, & Van

der Heijden, 1994). This notion that attention is guided to target features is at the heart of the Guided Search model of Wolfe, Cave, and Franzel (1989). Guided Search has been updated and modified over the years in several labs (Moran, Zehetleitner, Müller, & Usher, 2013; Schwarz & Miller, 2016; Wolfe, 1994, 2007; for a review, see Wolfe, 2014). Key to most versions is the idea that a limited set of attributes can guide search. This set starts with the "preattentive" features proposed by Treisman (1985, 1986), though it has been elaborated by much work in many labs over the subsequent years (reviewed in Wolfe & Horowitz, 2017).

In this paper, we ask about the time course of guidance. If a feature like the color of objects can guide search, how long does it take to get some guidance started and how long does it take for that guidance to reach its full strength once the feature is visible? Aspects of this and related questions have been addressed in previous work. There is an extensive literature showing that the brain can be prepared to give preference to stimuli showing specific features in humans (reviewed in Scolari, Ester, & Serences, 2014) and monkeys (reviewed in Treue, 2014). Thus, if the target is going to be red, cells that prefer red can be boosted across the entire visual field; a neural embodiment of guidance. Behaviorally, Olds and her colleagues (Olds, Cowan, & Jolicœur, 2000a, 2000b) performed

✉ Evan M. Palmer
   evan.palmer@sjsu.edu

[1] Visual Attention Lab, Brigham & Women's Hospital, Boston, MA, USA

[2] Departments of Radiology and Ophthalmology, Harvard Medical School, Boston, MA, USA

[3] Department of Psychology, San José State University, One Washington Square, San Jose, CA 95192-0120, USA

⩙ Springer

a series of experiments in which the time course of "pop-out" search was measured by adding distractors to a search task to disrupt efficient search. If observers are looking for a target of one color among distractors of another homogeneous color, search is very efficient; the target is said to "pop out." If there are distractors of two colors and the target color lies on a line in a color space between those two distractors, search is inefficient (Bauer, Jolicœur, & Cowan, 1996). Olds et al. presented the target and distractors of one color followed, after a stimulus onset asynchrony (SOA), by the addition of distractors of the other color. If the SOA was zero, this was an inefficient search. With even 50-ms SOA, there was some benefit provided by the period of one distractor color search. RTs continued to decline as the SOA was increased to 1,000 ms, with the bulk of the improvement seen within 200 ms. Further analysis led the authors to conclude that RTs were a mix of trials in which attention was efficiently guided to the color singleton and trials in which it was necessary to search randomly among items until the target was discovered. The data suggest that the color singleton could usually guide the search to the feature when given about 200 ms. The Olds paradigm addresses how long a feature search takes to complete. It does not directly ask how long it takes for the feature guidance to get going.

Olds and Fockler (2004) tried to measure the effects of pre-cuing by features in a conjunction search. Their observers were looking for targets defined by a conjunction of color and orientation. They provided pre-cues in which either the color information or the orientation information was specified first, with the full display presented after a delay. Thus, red and green squares might be presented initially, followed after a delay by the offset of parts of the red and green squares to turn them into vertical and horizontal elements. Though there were small effects, the preview period did not substantially change the subsequent search, perhaps because it did not provide any more guidance than was available in the zero SOA case when color and orientation information became available at the same time (see also Olds, Graham, & Jones, 2009; Sobel, Pickard, & Acklin, 2009).

Wolfe et al. (2004) cued the identity of a search target prior to the appearance of the search display in a series of experiments. In their Experiment 1, items could be big or small, red or green, and vertical or horizontal. On any given trial, there would be two types of distractor and, on 50% of the trials, one unique target. As a baseline condition, that target – e.g., red vertical – would remain consistent for an entire block of trials. The RT for this consistent condition was compared to conditions where the target changed on each trial. On those trials, observers got a verbal or picture cue of the target at some SOA prior to the appearance of the search display. Picture cues were more effective than word cues. Picture cues were fully effective within 200 ms. Word cues continued to become more effective over 800 ms. For both cues, very substantial effects were seen within 50–100 ms.

Note that the comparison in the Wolfe et al. (2004) article is to a consistent baseline condition. However, even in that consistent condition, there is a time course of interest. Even if you know that you are looking for "red," for example, you cannot start guiding to red *on this trial* until you see where the red items are located. The cueing experiments answer the question "How long does it take to convert information about target identity into effective guidance?" It does not quite answer the question "How long does it take guidance to get going?"

Other lines of experimentation have asked about how long it takes to begin to make use of information about the location of items. There is a vast literature on cues to target location, much of it rooted in the work of Posner (Posner, 1980; Posner & Cohen, 1984; Posner & Rothbart, 2007). Cues of various sorts can indicate the location of the target. This literature makes an important distinction between endogenous and exogenous cues (Klein, 2009). Valid exogenous cues are typically located at the location of the target. Endogenous cues are typically symbolic or linguistic cues, located at fixation; e.g., the word "left" to indicate a target at the left side target location. When time course has been studied (as reviewed in Luck & Vecera, 2002), exogenous cues have their effects within about 100 ms. By contrast, endogenous cues take longer (250–300 ms) to become fully effective. An invalid cue can still attract attention. After attention is directed elsewhere, the originally cued location will be inhibited. The inhibitory effect reaches its peak 300–400 ms after the onset of the invalid cue (Dukewich & Klein, 2015; Klein, 1988).

In a search for a conjunction of two features, it is possible to get a cueing benefit by informing the observer about items that will *not* be the target. This can be done by showing one of two sets of distractors before the other. Thus, in Watson and Humphrey's (1997) preview search experiments, observers might be looking for a blue H among green H and blue A distractors. If the set of green Hs is shown first, the observer somehow discounts those items, making the eventual search equivalent to a much more efficient search for a blue H among blue As. Observers may be inhibiting the old items, "visual marking" (Watson & Humphreys, 1997), prioritizing the new items (Donk & Theeuwes, 2003), or using the temporal offset to segregate new items from old (Jiang, Chun, & Marks, 2002). For present purposes, it is worth noting that preview search requires a significant amount of time. It takes about 400 ms of preview to effectively mark the old set in visual marking.

In the present experiments, we are trying to address a question that is somewhat different from the questions addressed in prior work. Suppose that you know something about what you are looking for. For instance, suppose you are looking for a valuable US penny. Unlike other US coins, pennies are copper colored, not silver. If a handful of coins is tossed before you, how quickly can you restrict your search to the copper-colored subset?

Experiment 1

The basic strategy of these experiments is cartooned in Fig. 1. Observers are asked to do an inefficient search for a Landolt C target (Vlaskamp, Over, & Hooge, 2005). One C is facing left or right and the observer needs to identify that direction. All other Cs open at top or bottom. In the absence of guidance, observers will need to deploy attention to item after item until they stumble upon the target (note that unguided search displays are depicted in the bottom of Fig. 3 but not depicted in Fig. 1). We can establish average RTs for this inefficient search for set sizes of four and 16. On guided trials, at some point, before or after the onset of the Cs, guiding color information will be made available, marking four of 16 items as possible targets. If guidance is perfect, the RT for a guided, set size 16 trial should be the same as the RT for the unguided, baseline set size 4. By measuring RT for a variety of SOAs, we can map the time course of guidance by mapping the transition from no guidance (RT equivalent to unguided 16) to full guidance (RT equivalent to unguided 4).

Simulation

To visualize the expected pattern of data from this experimental design, consider this simple simulation. On each trial of the simulation, there are 16 items. One of these is a target. Four items, including the target, are in the guiding subset (equivalent to the smaller subset of colors in the actual experiment). There is an SOA between the onset of the items and the onset

of the color. If the SOA is negative, items appear before the color. If the SOA is positive, colors appear before the items (as in Fig. 1). Guidance starts with some delay after the onset of the color. An unguided search is a search through 16 items. A guided search is a search through the four items of the correct color.

On each simulated deployment of attention, an item is selected at random from the effective set size; 16 if guidance is not in place, four if it is in place. We present two versions of the simulation. In Fig. 2a, items are selected *without replacement* assuming perfect "inhibition of return" (Klein, 2000). However, as can be seen in Fig. 2b, the basic pattern of results does not change if we assume sampling *with replacement* (Horowitz & Wolfe, 1998), assuming no inhibition of return (Klein, 2000). In this paper, we do not want to debate the question of memory for rejected distractors. We ran the simulation both ways in order to illustrate that, in this case at least, the question of memory for rejected distractors does not matter much. The RT for each trial is the number of deployments of attention needed before the target is found.

Results are shown in Fig. 2. Dashed lines show the average number of deployments required for unguided searches through set sizes four and 16. The curves show how RT changes as a function of SOA for three different guidance delays (all time measurements are in units of deployments of attention). Notice that the simulation predicts guidance benefits for negative SOAs; those cases where the color appears after the search items appear. This occurs because some searches that
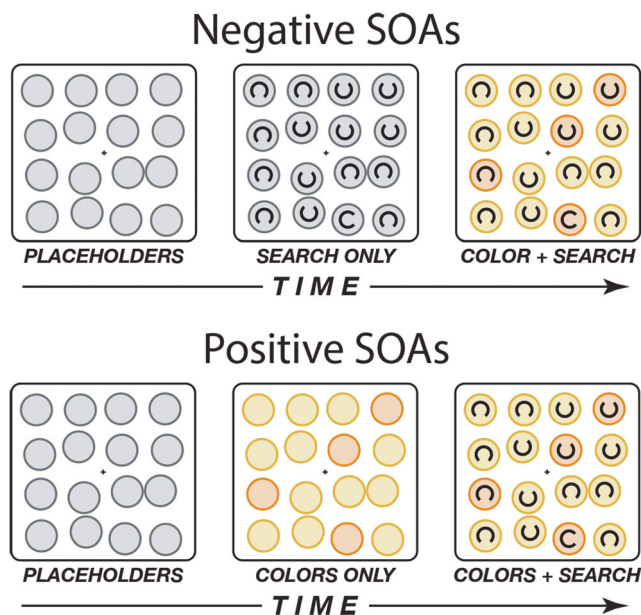
Fig. 1 Either the Landolt C search display appeared before the colored circles (top row, negative SOAs) or the colored circles appeared before the search display (bottom row, positive SOAs). A 0-ms SOA condition was also tested in which both the search display and the colored circles appeared at the same time
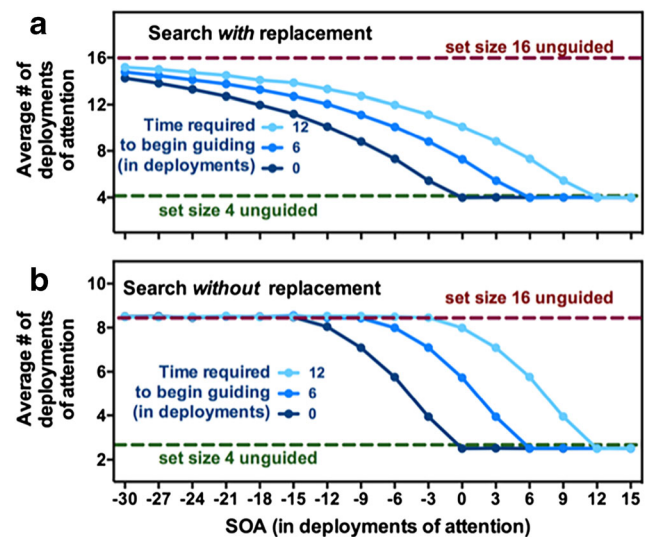
Fig. 2 Speed of guidance simulation. (a) Search with replacement. (b) Search without replacement. Times are expressed as number of attentional deployments. Dashed lines show unguided searches through four or 16 items. Solid lines show RT as a function of the SOA between the onset of search items and guiding colors (negative SOA indicates search items appear first). The parameter is the time to begin guidance. As SOA increases, search becomes more guided. Search is perfectly guided when the SOA is longer than the time required to begin guidance. Simulated data from 50,000 trials per data point

begin in an achromatic state last until after the colors have appeared and guidance has started. Unsurprisingly, as can be seen in the figure, these curves reach the perfectly guided set size 4 floor when the SOA is equal to the time required to begin guidance. This intercept can be used to derive an estimate of the time to guide from the actual data. Changes in the assumptions – e.g., assuming inhibition of return or making the transition from unguided to guided gradual, rather than abrupt – do not change the interpretation of the data. The estimate derived from the SOA intercept is the estimate of the time required to get guidance going.

Method

Participants Seven males and eight females, average age 32.13 (SD: 10.36) years, from the Brigham & Women's Hospital Visual Attention Laboratory's paid subject pool participated in the experiment and received US$10 per hour compensation. All observers provided informed consent, passed an Ishihara Color Test, had 20/25 or better vision (with correction, if necessary), and reported no history of eye disease or muscle disorders. In this and the following experiments, we included 13 or 15 participants since similar experiments in this field typically employ sample sizes of 10–15 participants (e.g., Wolfe, et al., 2004; Palmer, Horowitz, Torralba, & Wolfe, 2011; Wolfe & Van Wert, 2010). It is not immediately clear how one calculates power for an experiment where one is collecting many RTs at multiple SOAs. At a crude level, mean RTs have a standard deviation of about 400 ms. With 15 participants, we can detect a mean RT difference of about 400 ms (half the difference between set size 4 and set size 16 baselines) with a power of 0.8 and an alpha of 0.05, but this is a very conservative estimate.

Materials Stimuli were presented and responses gathered on an Apple Macintosh G4 450 MHz computer driving a 20-in. (diagonal) CRT monitor at a resolution of 1,024 × 768 pixels. Matlab 5.2.1 and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) were used to control the presentation and timing of the stimuli.

At the viewing distance of 57.4 cm, the display area was a square measuring 22.5 × 22.5° visual angle. This region was divided into an invisible 4 × 4 array of cells with each cell subtending approximately 5.63 × 5.63°. Circles containing the Landolt C stimuli were 3.75 × 3.75° and randomly jittered both vertically and horizontally by up to ± 1° from the center of each cell.

Procedure Participants were seated in a darkened room in front of the testing computer. Instructions were delivered on the computer screen as part of the testing protocol. Participants completed a practice block of 26 trials, which was not included in the analyses. Following the practice block,

participants completed nine blocks of 160 trials each, with time for a break between each block. Total testing time was approximately 1.5–2 h and was completed in a single session.

Design Participants searched displays consisting of four or 16 colored circles, each containing a Landolt C stimulus (i.e., an outline circle with a small gap; Figs. 1 and 3). All Cs had gaps on the top or bottom except for the target stimulus, which had a gap on either the left or right. Participants indicated as quickly as possible whether the target C's gap was on the left or the right by pressing either the <A> key (left gap) or <"> key (right gap). Targets were present on every trial.

Initially, all 16 circles in the display were gray (R: 219, G: 222, B: 226) placeholders that appeared on a white background (R: 255, G: 255, B: 255). During the trial, two sources of information were provided to the observer, and the timing of when this information replaced the placeholder stimuli was of critical interest for the study. One piece of information was the appearance of the Landolt C search stimuli. Once the Cs appeared, the observer could begin to search for the target. RTs for each trial were measured from the time the Landolt C stimuli appeared until the participant made a key press.

Another source of information was the color of the circles. At some point during the trial, the gray placeholder circles were replaced with colored circles. On the majority of trials, visual attention could be guided by color since 12 of the circles were one color and the remaining four circles were another color. Color combinations used were blue (R: 177, G: 209, B: 237) and green (R: 188, G: 221, B: 199), purple (R: 235, G: 214, B: 224) and red (R: 230, G: 192, B: 192), and orange (R: 243, G: 218, B: 192) and yellow (R: 242, G: 232, B: 197). Participants knew that the black (R: 35, G: 31, B: 32) target C would always reside in the smaller set of colored circles in these situations. (Note that colors used in this experiment were not equiluminant.)
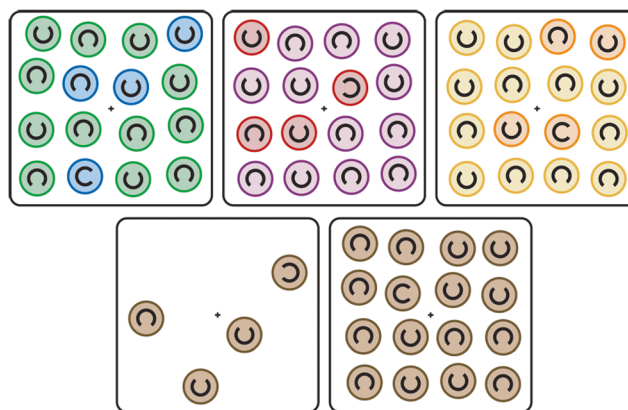

Fig. 3 Five sample displays illustrating the colored stimuli used in Experiment 1. In the top row displays support guided search since observers knew that the target resided in the smaller, four-item color subset. The bottom row of unguided search stimuli shows baseline conditions with set sizes four and 16 but no guiding color information

The remaining unguided baseline search displays had circles all of a single color, which provided no information to the observer about the location of the target. Baseline displays either contained four or 16 brown (R: 108, G: 87, B: 53) circles, representing "best-case" and "worse-case" baseline conditions for the search. Color in the baseline conditions was revealed in the same manner as the guidance conditions, except that participants only ever saw four or 16 Landolt Cs and they were all the same color (brown). Thus, for example, a set size 4 baseline condition with a negative SOA first had four gray placeholders appear, and then the Landolt C stimuli on a gray background, then the backgrounds turned brown. For a set size 4 baseline condition with a positive SOA, first four gray placeholders appeared, then the placeholders turned brown, then the Landolt Cs appeared. In other words, the timing of appearance of the background colors and Landolt Cs was the same as the guidance conditions, but all pedestals had the same color, meaning that no guiding information was available.

The SOA between the appearance of the colored circles and Landolt Cs was the critical manipulation in this study. As indicated in Fig. 1, negative SOAs indicate that the Landolt Cs were displayed before the colored circles, positive SOAs indicate that colored circles were presented before the Landolt C stimuli appeared. An SOA of 0 ms means that both the Landolt Cs and the colored circles appeared at exactly the same time. SOAs for the single-color baseline condition were: -800, -360, 0, 360, and 800. SOAs for the multiple color conditions were: -800, -600, -480, -360, -240, -120, 0, 120, 240, 360, 480, 600, and 800. Each baseline display and SOA combination was tested 27 times, spread randomly throughout the nine experimental blocks. The multiple color conditions were each tested 30 times per block.

Figure 4 depicts three types of block structure used in this experiment for guided search displays. Blocks of consistent, mixed, or interference conditions consisted of 160 trials, including 30 baseline search displays (four or 16 brown circles) randomly intermixed. Consistent blocks had the same color scheme for all guidance trials in a block (e.g., four red circles and 12 purple circles, on every two-color trial). In mixed blocks, the colors presented varied from trial to trial, but the assignment of color to either the large or small set of circles was consistent during the entire block (i.e., targets were always red among purple, blue among green, or orange among yellow). Interference blocks also had color sets that varied from trial to trial, but the assignment of color to the large or small set of circles reversed every time those colors appeared (i.e., red among purple, then purple among red, then red among purple, with other color pairs interspersed and also reversing upon every presentation).
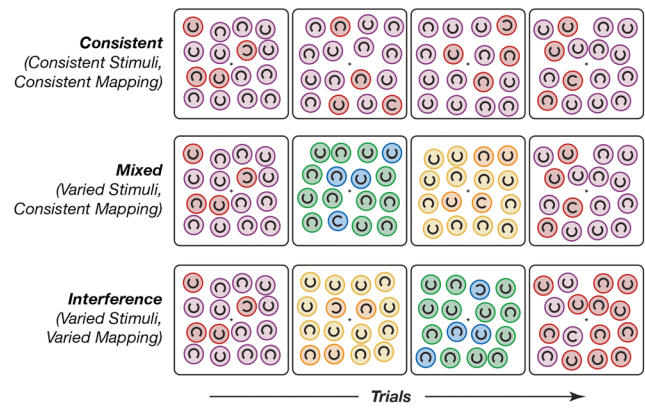
Fig. 4 Within a block of 160 trials, one of three types of structure was present for displays supporting guided search. In the consistent condition (top row), the same color stimuli were used on each trial and they were consistently mapped. In the mixed condition (middle row), colors varied randomly from trial to trial, but a target subset color (e.g., red) remained a target color for the whole block. In the interference condition (bottom row), colors varied randomly and a target color on one trial became the distractor color on its next appearance

Results and discussion

Data analysis methods Figure 5 shows mean RTs, averaged over participants for each combination of SOA, block type, and display type. The dense sampling of SOAs in the three
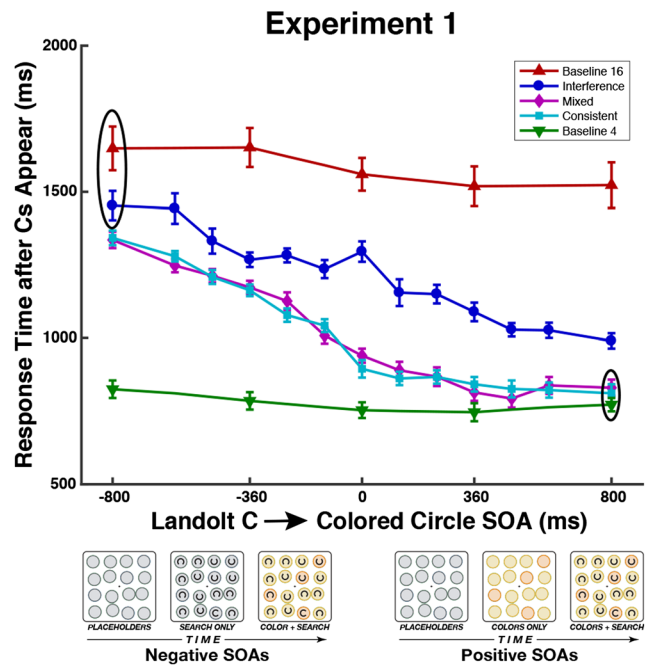
Fig. 5 Mean response time after Landolt C appearance as a function of Landolt C – Colored Circle SOA in Experiment 1. The three guidance conditions were tested against the one-color baseline 16 and one-color baseline 4 conditions at SOAs of -800, -360, 0, 360, and 800. Data points from the guidance conditions are significantly different than the baseline conditions unless circled. All data points within a circle have BF scores less than 3 and are statistically indistinguishable from the respective baseline conditions. Error bars indicate within-subjects confidence intervals (Cousineau, 2005; Morey, 2008)

guidance conditions allows us to visualize the whole SOA function. The primary messages to take from these data are that guidance develops gradually over many hundreds of milliseconds and that it is not fully available when the guiding signal and the Landolt Cs appear simultaneously (SOA 0 ms). Full guidance is available only when the guiding signal is made available hundreds of milliseconds before the onset of the Cs. At that point, guidance to a subset of four items looks like search through a set of four items.

Since the baseline conditions used a subset of the SOAs, we confined our formal statistical analysis to those conditions where data were collected for the three guidance and two baseline conditions (i.e., -800, -360, 0, 360, and 800 ms SOA). Significant effects from the ANOVA are examined with t-tests, employing a Bonferroni correction for multiple comparisons to help avoid type I errors. Since some of the theoretical arguments we are making rely on determining at which SOAs the guidance conditions differ (or not) from the baseline conditions, we also performed Bayes factor (BF) calculations with default priors (Morey & Rouder, 2015; Rouder, Speckman, Sun, Morey, & Overson, 2009) for every planned comparison using the JASP software package (JASP Team, 2018). Critically, such BF calculations provide relative evidence in favor of *or against* rejecting the null hypothesis, with BFs less than one providing evidence for accepting the null hypothesis (Jeffreys, 1961; Kass & Raftery, 1995). In visual depictions of the data set, within-subject confidence intervals were created for all data points using the Cousineau (2005) method with the Morey (2008) correction.

We conducted a 5 × 5 (Display × SOA) repeated measures ANOVA on the mean RTs from each participant in both the baseline conditions and the three guidance conditions with matching SOAs. Mauchley's test of sphericity revealed significant effects for both display, $\chi^2(9) = 64.39$, $p < .001$, and SOA, $\chi^2(9) = 21.26$, $p < .05$, therefore degrees of freedom were corrected using Greenhouse-Geisser values for sphericity ($\varepsilon = .32$ and $\varepsilon = .52$, respectively). The ANOVA revealed main effects of both display, $F(1.27, 17.88) = 37.67$, $p < .001$, $\eta_p^2 = .73$, and SOA, $F(2.07, 28.99) = 95.60$, $p < .001$, $\eta_p^2 = .87$, as well as a significant display × SOA interaction, $F(6.05, 84.64) = 16.53$, $p < .001$, $\eta_p^2 = .54$.

With regard to the main effect of display, planned comparisons of the mean RTs for the five display conditions (with a Bonferroni correction, alpha = .005) revealed that latencies in the interference condition were longer than in the mixed, $t(14) = 3.79$, $p < .005$, $d = .59$, two-tailed, BF = 21.56, or consistent conditions, $t(14) = 4.17$, $p < .001$, $d = .62$, two-tailed, BF = 41.00. The slower RTs in the interference condition probably reflect the negative effects of priming. It is known that the target features of the current trial prime those features for the next trial (Kristjansson, 2006; Kruijne & Meeter, 2015; Maljkovic & Nakayama, 1994). In the interference condition, those primed features can turn out to be the features of the

distractor on the next trial, weakening guidance and slowing search. The mixed and consistent conditions were not reliably different from each other using frequentist statistics, $t(14) = 0.36$, $p = .72$, and the BF score of 0.28 provides "positive" evidence for the null hypothesis under a Bayesian perspective (Jeffreys, 1961; Kass & Raftery, 1995).

With regard to the main effect of SOA, RTs generally decreased as SOA increased, presumably driven by more efficient guided search performance at longer SOAs in all three guidance conditions. Planned t-tests and BF calculations established that all pairwise differences between SOAs were reliably different, all $t(14) \geq 5.13$, all $p \leq .0002$, all $d \geq .30$, two-tailed, all BF ≥ 193.30 except for the comparison of +360 versus +800 ms SOA, $t(14) = 1.45$, $p = .17$, $d = .06$, two-tailed, BF = .64. The BF score indicates modest support for the null hypothesis model of no difference between the +360 and +800 ms SOA conditions (Kass & Raftery, 1995), by which point the RTs have reached an asymptote at roughly the level of the set size 4 baseline.

Our particular interest in this study is reflected in the significant interaction of display × SOA, in which we are focused on when the three guidance conditions diverged from the baseline 16 condition (start of guidance) and when they converged with the baseline 4 condition (guidance reaches maximum). These trends were examined through planned comparisons of the three guidance conditions against each of the baseline conditions (set sizes four and 16) using both t-tests and BF calculations, employing a Bonferroni correction to guard against type I error (alpha = .0017) for the t-tests. T-tests and BF results are presented in Table 1a and b, separately for the set size 4 and 16 baseline conditions. In Fig. 5, data points that are circled failed to register as significantly different from each other under frequentist statistics, all $p > .05$, and had BF scores ≤ 3, indicating differences "not worth more than a bare mention" (Kass & Raftery, 1995). Uncircled data points are significantly different from each other, all $p \leq .0017$, two-tailed, all BFs > 3.

Figure 5 shows guidance developing over several hundred milliseconds. Starting at the left, when the Landolt Cs appear before the guiding color information, unguided search through the Cs begins. The difference between each of the guided conditions and the baseline 16 average becomes significant for SOAs of -360 ms and later (Table 1a). Most of the value of color guidance is in place when color and Cs appear at the same time (SOA 0), though having the color appear first (positive SOA) does provide more guidance than is available at SOA 0. The consistent condition is statistically distinguishable from the baseline 4 condition until +360 ms SOA, and both the consistent and mixed conditions are statistically distinguishable from the baseline 4 condition until +800 ms SOA (Table 1b).

The finding that the consistent, mixed, and interference conditions were all significantly different than the baseline 4 condition at 0 ms SOA (with BF scores ≥ 1,497) means that participants were not able to make full use of guiding information in the

Table 1 A and B T-test t-values (top), p-values (middle, italicized), and Bayes factor values (bottom) from planned comparisons of the three guidance conditions against the baseline 16 (Table 2A) and baseline 4 (Table 2B) conditions in Experiment 1

Table 1A

		Baseline 16				
		-800	-360	0	360	800
Consistent	t-value:	3.23	4.85	6.55	6.16	5.31
	p-value:	0.0061	0.0003	0.0000	0.0000	0.0001
	BF:	8.39	124.69	1752.68	983.11	264.74
Mixed	t-value:	3.19	4.66	6.59	6.42	5.53
	p-value:	0.0066	0.0004	0.0000	0.0000	0.0001
	BF:	7.84	91.81	1863.80	1453.24	374.77
Interference	t-value:	2.77	5.28	3.81	5.88	4.67
	p-value:	0.0150	0.0001	0.0019	0.0000	0.0004
	BF:	4.00	252.83	22.39	646.31	93.14

Table 1B

		Baseline 4				
		-800	-360	0	360	800
Consistent	t-value:	10.11	11.55	6.44	3.11	1.64
	p-value:	0.0000	0.0000	0.0000	0.0076	0.1223
	BF:	169018.21	783378.19	1497.18	6.97	.79
Mixed	t-value:	10.40	10.74	7.47	4.63	2.19
	p-value:	0.0000	0.0000	0.0000	0.0003	0.0463
	BF:	235406.16	337367.56	6458.48	86.63	1.63
Interference	t-value:	6.92	8.97	8.15	4.80	5.95
	p-value:	0.0000	0.0000	0.0000	0.0003	0.0000
	BF:	3000.42	44965.75	15942.49	116.17	712.23

White squares indicate evidence in favor of the null hypothesis, gray squares indicate an effect difference "not worth more than a bare mention," light green squares indicate "positive" evidence of a difference, medium green squares indicate a "strong" difference, and dark green squares indicate a "very strong" difference (Kass & Raftery, 1995)

displays immediately on its appearance. In other words, a search set with four targets of one color and 12 distractors of another color is not searched as efficiently as a display with just four items, when all items are presented simultaneously. This effect held, even for the consistent mapping condition, in which participants saw the same color mappings (e.g., red among purple) for over a hundred trials within a single block.

It is possible that the baseline 4 condition gains some advantage because it is less crowded than 16-element displays. That effect would be roughly constant so its magnitude cannot be much bigger than the small difference between baseline 4 and subset conditions at large SOAs. Moreover, color differences tend to weaken crowding effects (Kooi et al., 1994; Fox; 1998), suggesting the small subset should be relatively protected from crowding effects.

It is worth noting that the guided conditions are faster than the baseline 16 condition, even at an SOA of -800 ms (color appears 800 ms after Cs.). This is not statistically reliable under our conservative Bonferroni correction but the BF evidence for a difference is "substantial" for the consistent and

mixed conditions albeit only "barely worth mentioning" for the interference condition (Jeffreys, 1961). These modest differences are suggestive and make theoretical sense. The RTs for an unguided search through 16 items average about 1,600 ms. That means that, even if the color appears 800 ms after the Cs and guidance is not complete for another few hundred milliseconds after that, there is still time left for guidance to help – and it probably does.

Taken together, these results suggest that color guidance is not fully effective immediately on color onset. It appears to be fully engaged after several hundred milliseconds of display time. This is consistent with other estimates, discussed in the introduction, that showed guidance developing over time before reaching full effectiveness.

Experiment 2 – more vivid colors

In Experiment 1, while color information was not completely effective immediately, it could be objected that the colors were

rather desaturated and similar in hue. Perhaps the apparent sluggishness of guidance was due the use of weak colors. Accordingly, Experiment 2 replicates Experiment 1 using stimuli with stronger color contrast. Figure 6 depicts the stimuli used in this study. Colors are more saturated and color pairs were highly dissimilar. Finally, the Landolt Cs were white, with a thin black outline to heighten contrast against the more saturated colors.

Methods

A sample of four males and nine females, average age 27.85 (SD: 9.45) years, from the Brigham & Women's Hospital Visual Attention Laboratory's paid subject pool participated in the experiment. Placeholder stimuli were gray (R: 128, G: 128, B: 128) and baseline stimuli were brown (R: 149, G: 121, B: 102). The vivid color pairs were orange (R:247, G:148, B:30) and blue (R: 51, G: 59, B: 151), cyan (R: 0, G: 183, B: 241) and red (R: 237, G: 28, B: 36), or purple (R: 143, G: 53, B: 140) and green (R: 0, G: 170, B: 79). Landolt Cs were white (R: 255, G: 255, B: 255) with a thin black outline (R: 35, G: 31, B: 32). Methods were otherwise identical to those in Experiment 1.

Results and discussion

Mean RTs as a function of SOA are shown in Fig. 7. A 5×5 (Display × SOA) repeated measures ANOVA was conducted on the mean RTs in both the baseline conditions and the three guidance conditions with matching SOAs. Mauchley's test of sphericity revealed significant effects for both display, $\chi^2(9)=$ 40.73, $p < .001$, and SOA, $\chi^2(9)= 27.65$, $p < .01$, therefore degrees of freedom were corrected using Greenhouse-Geisser values for sphericity ($\varepsilon = .42$ and $\varepsilon = .48$, respectively).

As in Experiment 1, the ANOVA revealed a main effect of display, $F(1.68,20.24) = 39.49$, $p < .001$, $\eta_p^2 = .77$, and both Bonferroni-corrected t-tests and BF calculations revealed that
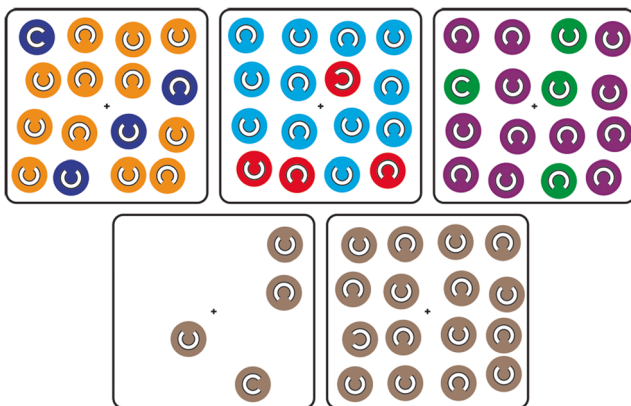
Fig. 6 Five sample displays from Experiment 2. Stimuli in this experiment had more saturated colors, stronger color contrast, and higher contrast Landolt Cs than the stimuli used in Experiment 1
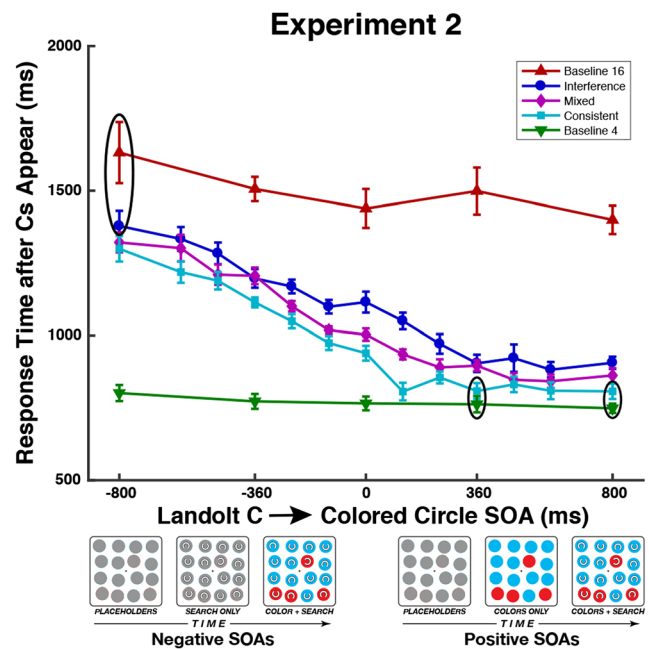
Fig. 7 Mean response time after Landolt C appearance as a function of Landolt C – Colored Circle SOA for 13 observers in Experiment 2. The three guidance conditions were tested against the one-color baseline 16 and baseline 4 conditions at SOAs of -800, -360, 0, 360, and 800. Data points from the guidance conditions are significantly different than the baseline conditions unless circled. All data points within a circle have BF scores less than 3 and are statistically indistinguishable from the respective baseline conditions. Error bars indicate within-subjects confidence intervals (Cousineau, 2005; Morey, 2008).

average RTs in the baseline 4 and baseline 16 conditions were significantly different than all other conditions, all $t(12) \geq$ 4.76, all $p \leq .0005$, all $d \geq 1.06$, two-tailed, all BFs ≥ 76.87. None of the three guidance display conditions (consistent, mixed, and interference) were different from each other, all $t(12) \leq 2.38$, all $p \geq .035$, all $d \leq .66$, two-tailed, all BFs ≤ 2.13. This means that, unlike Experiment 1, the interference condition did not yield significantly slower RTs in the aggregate than the consistent and mixed conditions. In fact, the BF analysis yielded some evidence in favor of the null hypothesis for the comparison of the mixed and interference conditions, $t(12)$ $= 1.00$, $p = .34$, $d = .28$, two-tailed, BF = 0.42.

The ANOVA also detected a main effect of SOA, $F(1.91,22.89) = 55.69$, $p < .001$, $\eta_p^2 = .82$, reflecting the fact that, overall, RTs decreased as SOA increased. Secondary analyses with t-tests and BF calculations revealed that all pair-wise combinations of SOAs were reliably different than each other, all $t(12) \geq 4.46$, all $p \leq .001$, all $d \geq .34$ two-tailed, all BFs ≥ 48.69 except for the comparison of SOAs +360 versus +800, $t(12) = 1.38$, $p = .19$, $d = .14$, two-tailed, BF = .61, which had some evidence in favor of the null hypothesis for that comparison.

The main effects were accompanied by a significant display × SOA interaction, $F(5.33,63.90) = 8.80$, $p < .001$, $\eta_p^2 = .42$, which was the primary statistic of interest for this study.

Planned comparisons examining the three guidance conditions against each of the baseline conditions (set sizes 4 and 16) employed a Bonferroni correction to guard against type I error (alpha = .0017). T-test and BF results are presented in Table 2a and b, separately for the set size 4 and 16 baseline conditions. In Fig. 7, data points that are circled failed to register as significantly different using both frequentist (all $p > .05$) and Bayesian (BF scores ≤ 3) statistical approaches, while uncircled data points are significantly different from each other under the traditional statistical approach (each $p ≤ .0017$, two-tailed), and had "positive evidence" for a difference (Kass & Raftery, 1995) under the Bayesian approach (each $BF ≥ 3$).

Importantly, all three guidance conditions were significantly different than the baseline 4 condition at 0 ms SOA, but were statistically indistinguishable by +360 ms SOA. Consistent with Experiment 1, color guidance took several hundred milliseconds to develop fully in displays with four targets and 12 distractors, to the point when

those displays were searched as quickly as displays with just four targets. By +360 ms, guidance appeared to be fully effective in the current study.

Overall, the major findings from Experiment 1 were replicated in Experiment 2. Participants were not able to fully use color guidance information immediately upon presentation (i.e., 0 ms SOA) but were able to by +360 ms SOA for the consistent display condition, despite the fact that the colors in this study were more saturated and salient than the colors used in the previous study. We take this to indicate the apparently slow time course of color guidance in Experiment 1 was not due to weak color signals. Stimuli with stronger color signals in Experiment 2 did not yield more rapid color guidance in this visual search task. For the consistent and mixed conditions at the -800 SOA value, it again seems that guidance was available toward the end of the average search, even when the colors appeared 800 ms after the Landolt Cs.

Table 2 A and B T-test t-values (top), p-values (middle, italicized), and Bayes factor values (bottom) from planned comparisons of the three guidance conditions against the baseline 16 (Table 2A) and baseline 4 (Table 2B) conditions in Experiment 2

Table 2A

		Baseline 16				
		-800	-360	0	360	800
Consistent	*t*-value:	3.18	4.65	5.75	6.24	7.18
	p-value:	*0.0080*	*0.0006*	*0.0001*	*0.0000*	*0.0000*
	BF:	6.98	64.78	309.75	604.04	1987.95
Mixed	*t*-value:	3.24	3.96	4.78	6.40	6.23
	p-value:	*0.0071*	*0.0019*	*0.0004*	*0.0000*	*0.0000*
	BF:	7.66	23.05	78.43	745.26	592.51
Interference	*t*-value:	2.43	4.17	2.90	5.83	5.61
	p-value:	*0.0316*	*0.0013*	*0.0134*	*0.0001*	*0.0001*
	BF:	2.30	31.64	4.56	348.90	255.56

Table 2B

		Baseline 4				
		-800	-360	0	360	800
Consistent	*t*-value:	8.53	12.45	5.53	1.61	2.17
	p-value:	*0.0000*	*0.0000*	*0.0001*	*0.1340*	*0.0510*
	BF:	9414.35	377505.58	228.58	0.78	1.59
Mixed	*t*-value:	13.16	6.91	9.24	3.53	3.34
	p-value:	*0.0000*	*0.0000*	*0.0000*	*0.0130*	*0.0059*
	BF:	661541.99	1416.77	20003.95	11.95	8.99
Interference	*t*-value:	8.24	7.01	6.49	2.91	3.50
	p-value:	*0.0000*	*0.0000*	*0.0000*	*0.0130*	*0.0044*
	BF:	6842.83	1602.96	834.41	4.68	11.44

White squares indicate evidence in favor of the null hypothesis, gray squares indicate an effect difference "not worth more than a bare mention," light green squares indicate "positive" evidence of a difference, medium green squares indicate a "strong" difference, and dark green squares indicate a "very strong" difference (Kass & Raftery, 1995)

Experiment 3 – colored Landolt Cs

The results of the first two experiments show that attentional selection is not immediately effective once guiding color information denoting the relevant subset of items with the target becomes available. However, in the previous two studies, the Landolt C stimuli were displayed on top of colored circles that provided the guiding information. Perhaps identifying the subset of circles containing the correct color information needed for guidance and then switching attention to the Landolt C stimuli inside took some extra time. That is, the color guidance information was near the target items but was not a part of the target items themselves.

To test this hypothesis, the stimuli in Experiment 3 were redesigned such that the target Landolt C stimuli were themselves colored (Fig. 8). We used the higher contrast colors introduced in Experiment 2. The stimuli all started out as complete circles and then a chunk of the circles disappeared at the appropriate time to turn them into Landolt C stimuli. Again, this sort of feature deletion from an object does not interrupt endogenous visual attention (Olds & Fockler, 2004; Theeuwes, 1991). If this experiment also shows that guidance takes time to develop, then we can be confident that the delay in the availability of guidance is not attributable to attention in having to switch between the colored circles used for guidance and the Landolt C target stimuli they contain. In this case, the Landolt C stimuli themselves contain the guiding color information so there needs to be no attentional switching between objects. One might ask why these stimuli were not used from the outset. These stimuli have the potential drawback that the color change would represent a change in the objects themselves, which might render them "new" objects at SOA 0. So
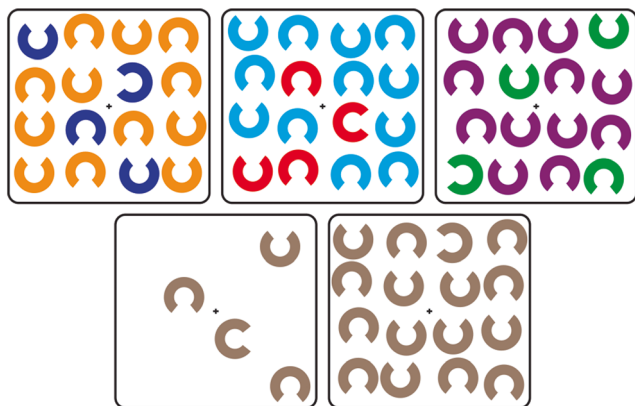
Fig. 8 Sample stimuli used in Experiment 3. The Landolt C stimuli were redesigned to be larger and incorporate the saturated colors introduced in Experiment 2. In this experiment, the guiding color information is part of the search stimuli themselves, potentially eliminating an extra attentional selection step

each stimulus set has advantages and disadvantages. However, as we show below, they all produce essentially the same results, giving us converging evidence for the conclusion that guidance takes time to develop after search stimuli become visible.

Methods

A sample of seven males and eight females, average age 28.80 (SD: 9.75) years, from the Brigham & Women's Hospital Visual Attention Laboratory's paid subject pool participated in the experiment. Colors used in Experiment 3 were the same as those used in Experiment 2; however, the Landolt C stimuli themselves were colored rather than being located within a colored circle (Fig. 8). Initially, the search stimuli were all gray outline circles (i.e., donut shaped). When the time came to reveal the Landolt C search stimuli, the appropriate chunk of each circle disappeared. This sort of stimulus offset does not interrupt endogenous visual attention (Theeuwes, 1991). Methods were otherwise identical to those in Experiments 1 and 2.

Results and discussion

Results are shown in Fig. 9 and are clearly similar to those of Experiments 1 and 2. Mean RT data in both the baseline conditions and the three guidance conditions with matching SOAs were submitted to a 5 × 5 (Display × SOA) repeated measures ANOVA. Mauchley's test of sphericity revealed significant effects for both display, $\chi^2(9)= 52.79$, $p < .001$, and SOA, $\chi^2(9)= 38.21$, $p < .01$, therefore degrees of freedom were corrected using Greenhouse-Geisser values for sphericity ($\varepsilon = .38$ and $\varepsilon = .41$, respectively). As in the previous studies, the ANOVA revealed main effects of both display, $F(1.53,21.35) = 34.54$, $p < .001$, $\eta_p^2 = .71$, and SOA, $F(1.64,23.00) = 49.74$, $p < .001$, , $\eta_p^2 = .78$, as well as a significant display × SOA interaction, $F(4.82,67.42) = 10.18$, $p < .001$, $\eta_p^2 = .42$.

The main effect of display was driven by all three of the guidance conditions (consistent, mixed, and interference) being significantly different than the baseline 16 condition, all ts(14) ≥ 4.89, all ps ≤ .0002, all ds ≥ 1.26, two-tailed, all BFs ≥ 134.80, and the baseline 4 condition, ts(14) ≥ 6.52, all ps ≤ .0001, all ds ≥ 1.27, two-tailed, all BFs ≥ 1673.50. There were no significant differences in overall RTs for the three guiding conditions in this experiment, all t(14) ≤ 2.74, all p ≥ .016, all d ≤ .41, all BFs ≤ 3.86, and actually the Bayesian analyses provided some evidence in support of accepting the null hypothesis for the comparison of mixed versus interference displays (BF = 0.56) and mixed versus consistent displays (BF = 0.85). As in Experiment 2 but not Experiment 1, the interference condition did not have significantly slower overall RTs than the consistent and mixed conditions.
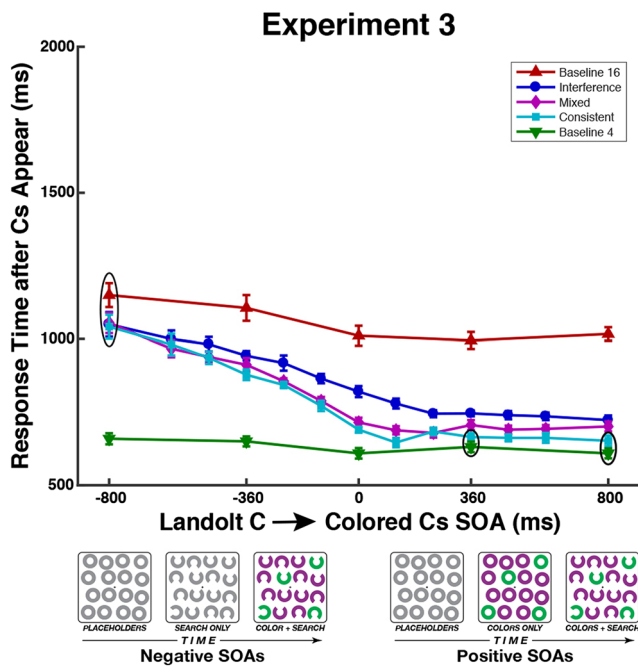
Experiment 3

Fig. 9 Mean response time after Landolt C appearance as a function of Landolt C – Colored C SOA for 15 observers in Experiment 3. The three guidance conditions were tested against the one-color baseline 16 and baseline 4 conditions at SOAs of -800, -360, 0, 360, and 800. Data points from the guidance conditions are reliably different than the baseline conditions unless circled. All data points within a circle have BF scores less than 3 and are statistically indistinguishable from the respective baseline conditions. Error bars indicate within-subjects confidence intervals (Cousineau, 2005; Morey, 2008).

To interpret the main effect of SOA, we used Bonferroni-corrected t-tests and BF analyses to evaluate all pairwise comparisons. These planned comparisons revealed that the -800 and -360 SOA conditions were significantly different than all other SOA conditions, all $t(14) \geq 4.19$, all $p \leq .001$, all $d \geq 1.08$, two-tailed, all BFs ≥ 42.31, but the +360 and +800 ms SOA conditions were not significantly different than the 0 SOA condition or each other, all $t(14) \leq 2.08$, all $p \geq .058$, all $d \leq .19$, two-tailed, all BFs ≤ 1.40. Furthermore, the Bayesian analyses provided evidence in favor of the null hypothesis for the 0 versus +360 (BF = 0.89) and +360 versus +800 (BF = 0.35) comparisons.

To better understand the significant display × SOA interaction, BF calculations and t-tests compared the three guidance conditions of the baseline conditions against each other (set sizes 4 and 16), employing a Bonferroni correction to guard against type I error (alpha = .0017). These planned comparisons are presented in Table 3a and b, separately for the set size 4 and 16 baseline conditions. The three guidance conditions become significantly different than the baseline 16 condition by -360 ms SOA for the consistent condition and 0 ms SOA for all conditions. The consistent condition does not become statistically indistinguishable from the baseline 4 condition until +360 ms SOA while the mixed and interference

conditions never become indistinguishable from baseline 4 condition under a Bayesian perspective, even at +800 ms SOA. As in the previous two experiments, guidance is not fully effective immediately (0 ms SOA) and takes time to develop (by +360 ms SOA for the consistent). In Fig. 9, circled data points are those that failed to register as significantly different from the baseline conditions under the frequentist statistical approach (all $p > .05$) and have a BF score of 3 or less from the Bayesian perspective. Uncircled data points are significantly different from baseline, each $p \leq .0017$, each $d \geq .69$, and each BF > 3, which indicates "positive evidence" of a difference (Kass & Raftery, 1995). Note that, in this experiment, there is less evidence of any guidance at an SOA of -800. Note also that the average RT is markedly faster in the 16 baseline condition of Experiment 3 than in the same condition in Experiments 1 and 2. Here, when the colors appear 800 ms after the Cs, the search is almost done and guidance cannot help very much, especially if it takes another 200–400 ms to reach full strength.

The results of this study are largely consistent with the previous two studies. Fully effective color guidance is not available at 0 ms SOA. It develops over several hundred milliseconds, becoming fully effective by +360 ms SOA for the consistent display condition. In the stimuli used in this experiment, color is a property of the Cs, not of the background. Therefore, the SOA functions of Experiments 1 and 2 were not a side effect of the need to switch attention between objects and their backgrounds.

General discussion

Looking at any of the data figures in this paper, we can narrate the time course of a search for the Landolt C, placed in the smaller, four-item color subset in a set of 16 items. If a display of four red and 12 blue items appears on the screen, it is quite clear from other work that the existence of that subset is available very rapidly. Work on ensemble processing (Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2003), rapid scene processing (Brady, Shafer-Skelton, & Alvarez, 2017; Greene, 2013; VanRullen & Thorpe, 2001), or enumeration (Watson & Maylor, 2006) tell us that information about the color of sets of items is available within the first 100 ms. However, the first deployments of attention are imperfectly guided by color. If the guidance were perfect, the RTs for finding the target C in the subset of four items would be the same as the RT for finding the target C when only four items are on the screen.

Given the RTs for the baseline 4 and baseline 16 conditions, we can calculate the cost of each additional item in this Landolt C task by calculating the slope of the RT × set size function. For Experiments 1 and 2, it is approximately 67 ms/item. For Experiment 3, it is about 40 ms/item. In Experiments 1 and 2, the difference between the mixed and consistent conditions and

Table 3 A and B T-test t-values (top), p-values (middle, italicized), and Bayes factor values (bottom) from planned comparisons of the three guidance conditions against the baseline 16 (Table 3A) and baseline 4 (Table 3B) conditions in Experiment 3

Table 3A

		Baseline 16				
		-800	-360	0	360	800
Consistent	t-value:	1.93	3.89	6.12	6.84	8.08
	p-value:	*0.0745*	*0.0016*	*0.0000*	*0.0000*	*0.0000*
	BF:	1.13	25.55	928.83	2663.82	14730.34
Mixed	t-value:	1.68	2.72	4.95	6.28	7.50
	p-value:	*0.1139*	*0.0167*	*0.0002*	*0.0000*	*0.0000*
	BF:	.83	3.66	147.98	1178.52	6646.11
Interference	t-value:	1.58	2.73	4.42	5.97	8.42
	p-value:	*0.1367*	*0.0163*	*0.0006*	*0.0000*	*0.0000*
	BF:	.73	3.74	61.62	733.62	23010.53

Table 3B

		Baseline 4				
		-800	-360	0	360	800
Consistent	t-value:	6.56	10.05	4.23	1.39	2.01
	p-value:	*0.0000*	*0.0000*	*0.0009*	*0.1868*	*0.0645*
	BF:	1788.23	163879.44	44.08	.59	1.26
Mixed	t-value:	8.86	11.14	8.05	3.59	4.71
	p-value:	*0.0000*	*0.0000*	*0.0000*	*0.0104*	*0.0003*
	BF:	39040.56	515928.49	14323.80	15.69	98.67
Interference	t-value:	5.30	9.44	6.30	2.95	3.96
	p-value:	*0.0001*	*0.0000*	*0.0000*	*0.0104*	*0.0014*
	BF:	259.62	80363.09	1204.81	5.38	28.61

White squares indicate evidence in favor of the null hypothesis, gray squares indicate an effect difference "not worth more than a bare mention," light green squares indicate "positive" evidence of a difference, medium green squares indicate a "strong" difference, and dark green squares indicate a "very strong" difference (Kass & Raftery, 1995)

the baseline 4 condition is about 200 ms at an SOA of 0 ms. This is the equivalent of adding 200/67 = 3 items to the set size. For Experiment 3, the difference between the mixed and consistent conditions and the baseline 4 condition is about 100 ms. This is the equivalent of adding 100/40 = 2.5 items. Thus, we can say that, when color and form appear at the same time, a set of four red and 12 blue items would behave like a set of six or seven red items. In the simulation shown earlier, this effective set size would arise because the first few selections of items would be randomly picked from all 16 items, leaving only a 25% chance that they would be drawn from the correct subset of four. Only after guidance became fully effective would attention be consistently directed to the appropriate subset of items. By 360 ms after the onset of the color, observers appear to be able to restrict search to that subset.

The longer a search goes on, the more effectively it will be guided to the correct subset of items. Thus, if the Cs come on, for example, 800 ms before the colors (SOA –800), the fastest searches will be unguided. The observer will simply have been lucky in his or her initial selection of Cs. After the colors

appear, the search will become increasingly guided so that we can assume that the last few selections of a long RT search will be almost entirely restricted to the correct color subset. This is seen more clearly in Experiments 1 and 2 because those tasks are harder, producing more long RTs that allow for the full development of guidance during the course of the search.

Guidance in these experiments can be seen as a cue that points to several locations at the same time. Cues that point to single locations (or objects) come in two forms, endogenous and exogenous (Klein, 2009). Exogenous cues like light onsets at the location of the target operate very quickly (~50 ms; Cheal & Lyon, 1991). Endogenous cues (e.g., semantic cues at fixation) take longer – perhaps, several hundred milliseconds (Horowitz et al., 2009; Nakayama & Mackeben, 1989).

While a color subset might be thought to be like an exogenous cue, a set of markers at the possible locations of the target, color guidance has a time course more like an endogenous cue. Interestingly, our lab has used various cues in search that are like classic endogenous cues including pictures of the target, presented at fixation. Generalizing over many experiments, these

seemingly endogenous cues seem to take no longer to reach their full potency than does the seemingly exogenous cue of color guidance (Wolfe, et al., 2004). In fact, it is probably a mistake to think of the color subset as *either* an endogenous or an exogenous cue. Cues direct focal attention to a location, guidance biases the deployment of attention. In the case of a color subset, that bias can be distributed to multiple locations in the display. It is probably more accurate to say that guidance happens to have a time course like an endogenous cue rather than to propose that they share an underlying mechanism, but this issue would require further investigation.

The current data indicate that guidance is not fully effective immediately at 0 ms SOA but is in place by +360 ms SOA. Why does it take as long as it does to get guidance fully engaged? Especially when the color of the subset is known, guidance could be implemented as a filter sitting athwart the feed-forward flow of information through the visual pathways. We know that this feed forward pathway is capable of accomplishing complicated acts of scene and object recognition in 100 ms or less (Fei-Fei, Iyer, Koch, & Perona, 2007; Serre, Oliva, & Poggio, 2007). There is also electrophysiological evidence for some color-based attention becoming effective within 100 ms of stimulus onset (Zhang & Luck, 2009). Other studies show signs of guidance that appear at about 250 ms (e.g., Berggren & Eimer, 2016). It is possible that the full implementation of guidance requires feedback from high cortical areas, as proposed in models like the Reverse Hierarchy Theory of Ahissar and Hochstein (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002) or the reentrant account of Di Lollo et al. (2000). We could hypothesize that guidance has a fast, feed-forward component and a slower component that requires feedback, but clear evidence for such a hypothesis awaits further research.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in the Cognitive Sciences, 8*(10), 457-464.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in the Cognitive Sciences, 15*(3), 122-131. https://doi.org/10.1016/j.tics.2011.01.003

Ariely, D. (2001). Seeing Sets: Representation by statistical properties. *Psychological Science, 12*(2), 157-162.

Bauer, B., Jolicœur, P., & Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly-separable from distractors. *Vision Research, 36*(10), 1439-1466.

Berggren, N., & Eimer, M. (2016). The control of attentional target selection in a colour/colour conjunction task. *Attention, Perception, & Psychophysics, 78*(8), 2383-2396. https://doi.org/10.3758/s13414-016-1168-6

Brady, T. F., Shafer-Skelton, A., & Alvarez, G. A. (2017). Global ensemble texture representations are critical to rapid scene perception. *Journal of Experimental Psychology: Human Perception and Performance, 43*(6), 1160-1176. https://doi.org/10.1037/xhp0000399

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

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *The Quarterly Journal of Experimental Psychology, 43*(4), 859-880.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*(4), 393-404.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*(1), 4-45.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*(4), 481-507.

Donk, M., & Theeuwes, J. (2003). Prioritizing selection of new elements: bottom-up versus top-down control. *Perception and Psychophysics, 65*(8), 1231-1242.

Dukewich, K., & Klein, R. (2015). Inhibition of return: A phenomenon in search of a definition and a theoretical framework. *Attention, Perception, & Psychophysics, 77*(5), 1647-1658. https://doi.org/10.3758/s13414-015-0835-3

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 32-39.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision, 7*(1), 10.

Fox, E. (1998). Perceptual grouping and visual selective attention. *Perception and Psychophysics, 60*(6), 1004-1021.

Greene, M. R. (2013). Statistics of high-level ccene context. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00777

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36*, 791-804.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*, 575-577.

Horowitz, T. S., Wolfe, J. M., Alvarez, G. A., Cohen, M. A., & Kuzmova, Y. I. (2009). The speed of free will. *The Quarterly Journal of Experimental Psychology, 62*(11), 2262-2288.

JASP Team (2018). JASP (Version 0.8.6) [Computer software].

Jeffreys, H. (1961). The Theory of Probability (3rd). Oxford University Press.

Jiang, Y., Chun, M. M., & Marks, L. E. (2002). Visual marking: Selective attention to asynchronous temporal groups. *Journal of Experimental Psychology: Human Perception and Performance, 28*(3), 717-730.

Kaptein, N. A., Theeuwes, J., & Van der Heijden, A. H. C. (1994). 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*(5), 1053-1069.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.

Klein, R. M. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*, 430-431.

Klein, R. M. (2000). Inhibition of return. *Trends in the Cognitive Sciences, 4*(4), 138-147.

Klein, R. M. (2009). On the Control of Attention *Canadian Journal of Experimental Psychology, 63*(3), 240-252.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*(2), 255-279.

Kristjansson, A. (2006). Simultaneous priming along multiple feature dimensions in a visual search task. *Vision Research, 46*(16), 2554-2570.

Kruijne, W., & Meeter, M. (2015). The Long and the Short of Priming in Visual Search. *Attention, Perception & Psychophysics, 77*(5), 1558-1573.

Luck, S. J., & Vecera, S. P. (2002). Attention. In H. Pashler & S. Yantis (Eds.), Stevens' Handbook of Experimental Psychology (3rd ed., Vol. 1: Sensation and Perception), pp. 235-286. New York: Wiley.

Maljkovic, V., & Nakayama, K. (1994). Priming of popout: I. Role of features. *Memory & Cognition, 22*(6), 657-672.

Moran, R., Zehetleitner, M., Müller, H. J., & Usher, M. (2013). Competitive guided search: Meeting the challenge of benchmark RT distributions. *Journal of Vision*, *13*(8), 24-24.

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*(2), 61-64.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor (Version 0.9.11-3)[Computer software].

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*(11), 1631-1647.

Olds, E. S., Cowan, W. B., & Jolicoeur, P. (2000a). The time-course of pop-out search. *Vision Research, 40*(8), 891-912.

Olds, E. S., Cowan, W. B., & Jolicoeur, P. (2000b). Tracking visual search over space and time. *Psychonomics Bulletin and Review, 7*(2), 292-300.

Olds, E. S., & Fockler, K. A. (2004). Does previewing one stimulus feature help conjunction search? *Perception, 33*(2), 195-216.

Olds, E. S., Graham, T. J., & Jones, J. A. (2009). Feature head-start: Conjunction search following progressive feature disclosure. *Vision Research, 49*(11), 1428-1447.

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search?. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 58.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Posner, M. I., & Cohen, Y. (1984). Components of attention. In H. Bouma & D. G. Bouwhuis (Eds.), Attention and Performance X (pp. 55-66). Hillside: Erlbaum.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58*, 1-23. https://doi.org/10.1146/annurev.psych.58.110405.085516

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237.

Schwarz, W., & Miller, J. (2016). GSDT: An integrative model of visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1654.

Scolari, M., Ester, E. F., & Serences, J. T. (2014). Feature- and Object-Based Attentional Modulation in the Human Visual System. In A. C. Nobre & S. Kastner (Eds.), Oxford Handbook of Attention (pp. 573-600). New York: Oxford U Press.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, *104*(15), 6424-6429.

Sobel, K. V., Pickard, M. D., & Acklin, W. T. (2009). Using feature preview to investigate the roles of top-down and bottom-up processing in conjunction search. *Acta Psychologica, 132*(1), 22-30.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Attention, Perception, & Psychophysics, 49*(1), 83-90.

Treisman, A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing., 31*, 156-177.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American, 255*, 114B-125.

Treue, S. (2014). Object- and feature-based attention: monkey physiology. In A. C. Nobre & S. Kastner (Eds.), Oxford Handbook of Attention (pp. 573-600). New York: Oxford U Press.

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30*(6), 655-668.

Vlaskamp, B. N. S., Over, E. A. B., & Hooge, I. T. C. (2005). Saccadic search performance: The effect of element spacing. *Experimental Brain Research, 167*(2), 246-259. https://doi.org/10.1007/s00221-005-0032-z

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review, 104*(1), 90-122.

Watson, D. G., & Maylor, E. A. (2006). Effects of color heterogeneity on subitization. *Perception & Psychophysics, 68*(2), 319-326.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*(2), 202-238.

Wolfe, J. M. (2007). Guided Search 4.0: Current Progress with a model of visual search. In W. Gray (Ed.), Integrated Models of Cognitive Systems (pp. 99-119). New York: Oxford.

Wolfe, J. M. (2014). Approaches to Visual Search: Feature Integration Theory and Guided Search. In A. C. Nobre & S. Kastner (Eds.), Oxford Handbook of Attention (pp. 11-55). New York: Oxford U Press.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1*, 0058. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., Horowitz, T. S., Kenner, N. M., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*(12), 1411-1426.

Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology*, *20*(2), 121-124.

Zhang, W. W., & Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nature Neuroscience, 12*(1), 24-25. https://doi.org/10.1038/nn.2223