Conjunction search: Can we simultaneously bias attention to features and relations?

Stefanie I. Becker[1] · Marina Atalla[1] · Charles L. Folk[2]

Abstract

Attention allows selection of sought-after objects by tuning attention in a top-down manner to task-relevant features. Among other possible search modes, attention can be tuned to the exact feature values of a target (e.g., red, large), or to the relative target feature (e.g., reddest, largest item), in which case selection is context dependent. The present study tested whether we can tune attention simultaneously to a specific feature value (e.g., specific size) and a relative target feature (e.g., relative color) of a conjunction target, using a variant of the spatial cueing paradigm. Tuning to the specific feature of the target was encouraged by randomly presenting the conjunction target in a varying context of nontarget items, and feature-specific versus relational tuning was assessed by briefly presenting conjunction cues that either matched or mismatched the relative versus physical features of the target. The results showed that attention could be biased to the specific size and the relative color of the conjunction target or vice versa. These results suggest the existence of local and relatively low-level attentional control mechanisms that operate independently of each other in separate feature dimensions (color, size) to choose the best search strategy in line with current top-down goals.

Keywords Attentional capture · Cognitive and attentional control · Cognitive control and automaticity

Conjunction search: Can we simultaneously bias attention to features and relations?

It is well known that we cannot process and consciously perceive all information present in a complex visual scene. Attention selects information for further in-depth processing.

Significance statement How do we search for a conjunction stimulus that consists of two features? Previous studies have shown that we preferentially tune attention to the relative features of a conjunction target (e.g., redder, larger) rather than its specific features (e.g., orange, medium). The present study shows, for the first time, that it is possible to tune attention to the relative feature of the target in one dimension (e.g., color), and the specific feature value in another dimension (e.g., size). This suggests the existence of local top-down attentional control mechanisms that adjust the search settings independently of each other within each stimulus dimension.

✉ Stefanie I. Becker
  s.becker@psy.uq.edu.au

[1] School of Psychology, The University of Queensland, McElwain Building, St Lucia, QLD 4072, Australia

[2] Department of Psychology, Villanova University, Pennsylvania, PA, USA

Hence, it is important to understand what factors control and guide visual attention. The currently prevalent view is that attention can be guided both by bottom-up, stimulus-driven factors (e.g., Theeuwes, 1992) as well as by the top-down goals of the observer (e.g., Folk & Remington, 1998; Wolfe, 1994; Yantis, 1993). For instance, perceptually salient stimuli such as items with a high feature contrast can reflexively attract or "capture" attention (e.g., Itti & Koch, 2001; Theeuwes, 1992). Attentional capture can, however, also be elicited by items that carry properties matching a sought-after item (e.g., Folk, Remington, & Johnston, 1992). For instance, when looking for a man in a red shirt, other red items such as a red umbrella, may involuntarily attract our attention (top-down contingent capture; e.g., Folk & Remington, 1998; Folk et al., 1992).

Top-down contingent capture is due to our ability to *tune* attention to the elementary feature(s) of the target, and initially it was believed that attention is usually tuned to the specific feature value of the target (e.g., specific color, size, orientation; e.g., Duncan & Humphreys, 1989; Folk & Remington, 1998; Martinez-Trujillo & Treue, 2004; Wolfe, 1994). In line with this contention, multiple visual search studies have shown that an irrelevant distractor that matches the target

feature (e.g., red) attracts attention much more strongly than other salient items that do not match the target (e.g., green; e.g., Folk & Remington, 1998). This similarity effect has typically been taken as evidence that attention was biased to the specific feature value of the target (e.g., Duncan & Humphreys, 1989; Ludwig & Gilchrist, 2002).

Subsequently, it was proposed that attention is not always tuned to the exact feature value of the target, but to an *optimal* feature value that best allows discriminating the target from the irrelevant items (e.g., Navalpakkam & Itti, 2007; see also Wolfe, 1994). According to the optimal tuning account (Navalpakkam & Itti, 2007; Scolari & Serences, 2009, 2012), especially when the nontargets are very similar to the target, it can be more optimal to tune attention to a feature value that is slightly shifted away from the nontargets to reduce activation by the nontargets and achieve a better signal-to-noise ratio.

Subsequent studies, however, have cast doubt on view that attention is usually tuned to the target's feature value or an optimal feature value (e.g., Becker, 2010). Rather, in line with a relational account, it has been shown that attention is usually biased to feature relationships, such as the relative color or size of the target (e.g., largest, reddest, darkest; Becker, 2010). For instance, in search for an orange target, attention would be tuned to all redder items, or *the reddest* item, when it is known or expected to be among the reddest items, and to all yellower or *the yellowest* item when the target is expected to differ in this direction from the irrelevant items. Contrary to feature similarity accounts or optimal tuning accounts, capture by an irrelevant item can depend only on whether it matches the relative feature of the target, independently of whether the feature values of the target and irrelevant item match (e.g., Becker, Folk & Remington, 2013).

For instance, in a spatial cuing study (Becker et al., 2013), participants had to search for an orange target among three yellow-orange nontargets, which should bias attention towards selection of the *reddest* item. The target was preceded by a salient, nonpredictive *cue* that was presented among three other cues that all had the same color (*cue context*). The colors of the cue and the cue context were varied so that the cue either had the same color as the target (orange), a different color (e.g., red or yellow-orange), and the same or a different *relative* color as the target (i.e., cue redder vs. yellower than the cue context). The results showed that a target-dissimilar, yellow-orange cue that had the same color as the nontargets still attracted attention when it was embedded in a yellow cue context and matched the targets' relative color (redder). A red cue that was embedded in a context of orange other cues that had the target color also captured attention, to the same extent, even though it competed with other cues that had the target color. In turn, a target-similar orange cue failed to attract

attention when it was embedded in a red cue context (yellower cue) and thus failed to match the relative target color (redder; Becker et al., 2013).

These results demonstrated that selection can be independent of the physical colors of target and cue or their similarity, and depend only on their relative match. This provides strong support for the relational account, and rules out current feature-specific views, including the feature similarity view and the optimal tuning account (Navalpakkam & Itti, 2007; Scolari & Serences, 2012; see also Geng, DiQuattro & Helm, 2017; Yu & Geng, 2019).

Subsequent studies extended these findings to conjunction search: Using a similar spatial cueing task with a color × size conjunction target and conjunction cues, Becker et al. (2017) showed that attention is tuned to the relative color and the relative size of a conjunction target. In the study, participants had to search for a blueish, large target among two blueish, medium and two bluish-green, large nontargets (bluer/larger target as the target was one of the bluest, largest items). A bluish-green, medium conjunction cue that was composed of the nontarget features, but had the relative color and size of the target (i.e., was presented among two bluish-green, small and two green, medium context cues) attracted attention, to the same extent as a physically matching control cue that had the same color and size as the target, and was embedded among context cues that matched the nontargets (Becker et al., 2017). Two other relatively matching cues that matched either set of nontarget features (bluish, medium conjunction cue and bluish-green, large conjunction cue) also attracted attention when they had the same relative features as the target, (larger/bluer), whereas physically matching cues that failed to match the relative feature(s) of the target failed to attract attention. The results were obtained behaviorally, as well as by measuring the N2pc in the EEG of participants, (a marker of visual attention; e.g., Eimer, 1996), and show that conjunction search is relational (in the sense that attention is tuned to the relative color and size of a conjunction target; Becker et al., 2017; see also Schönhammer, Grubert, Kerzel, & Becker, 2016). As capture was found for conjunction cues that matched the nontarget features, the results rule out an optimal tuning account or a feature similarity account. Moreover, since the conjunction cues and targets in the study did not have a unique feature, but differed only in a conjunction of features from other competing stimuli (nontargets and cue context), it is clear that attentional capture was not mediated by bottom-up feature contrasts or saliency, but was due to the manner in which attention was tuned to the target. As capture was only observed by cues that matched *both* relative features of the target, attention

was apparently tuned to both relative features of the target simultaneously (whereby we take "simultaneously" to mean "within the same trial" or "within the same selection episode"; Becker et al., 2017).

In addition, this and several other studies established that relational search seems to be a *default* search strategy (Becker, 2010; Becker et al., 2013; Becker, Harris, Venini, & Retell, 2014b), in the sense that it is preferred even when the target can theoretically be localized equally well by its singleton status (e.g., singleton detection mode; Bacon & Egeth, 1994) or its specific feature value (Brand, Oriet, Johnston, & Wolfe, 2014).

However, in line with current feature-specific accounts, it has also been shown that attention <u>can</u> still be tuned to the exact target feature value, at least in pop-out search and feature search tasks. In practice, feature-specific tuning to the exact target feature could only be observed when a relational search strategy had been prevented; for instance, when the features of the nontargets were varied such that the target was not reliably the largest or smallest item (or the reddest or yellowest item) anymore (e.g., Becker, Harris, Venini, & Retell,2014; Harris, Remington & Becker, 2013). However, compared with relational search, feature-specific tuning can result in less efficient search (e.g., Becker, Harris, Venini, & Retell, 2014b), and these differences in search efficiency between relational versus feature-specific search strategies may also (at least in part) explain the linear separability effect (D'Zmura, 1991)—that search is more efficient when the target has a relatively extreme feature value in feature space (linearly separable target; e.g., largest/smallest; steepest/flattest; darkest/brightest) than when it has an intermediate feature value (nonlinearly separable target; e.g., medium target among small and large nontargets; tilted target among vertical and horizontal nontargets; Bauer, Jolicoeur, & Cowan, 1996; Becker, 2010; Brand et al., 2014; Hodsoll & Humphreys, 2001).

In sum, previous research has shown that observers can adopt both relational and feature-specific search strategies (at least in pop-out and feature search tasks), whereby relational search seems to be the preferred search mode that also yields potentially more efficient search. However, it is currently unclear how these search modes are instantiated and, specifically, at what level of the visual system it is decided whether to engage in relational search or feature-specific search.

To date, there is evidence that top-down control can be implemented by both higher-order cognitive processes that operate on higher-level features (e.g., holistic target representations; Kahneman, Treisman, & Burkell, 1983; Treisman, Kahneman, & Burkell, 1983; or semantic properties; Koivisto & Revonsuo, 2007; Vecera & Farah, 1994; Wolfe & Bennett, 1997; Wyble, Folk, & Potter, 2013), as well as more local, low-level processes that adjust search settings semiautomatically on the basis of feedback—for instance, about the item selected on the previous trial (e.g.,

Becker, 2008a, 2008b; Irons & Leber, 2016; Maljkovic & Nakayama, 1994; Meeter & Olivers, 2014). However, little is known about the mechanisms that implement relational versus feature-specific top-down settings or their flexibility in adapting to different conditions across different stimulus dimensions.

Previous studies suggest that observers will usually start out with relational search and switch to feature-specific search only when the target cannot be found on the basis of relational search (e.g., Becker, Harris, Venini, & Retell, 2014b). That is, feature-specific search may be an adaptation or a change in search mode that occurs whenever relational search leads to frequent selection of nontargets. Such an adaptation or fine-tuning of the top-down setting could be made in response to on-line feedback, and could predominantly involve local processes that adjust the fine-tuning within a given stimulus dimension (e.g., only in the color dimension).

An alternative possibility is that the mechanism responsible for determining the search mode (relational or feature specific) operates at a more global level, and assigns a single search mode uniformly across all stimulus dimensions. This could be the case, for instance, if information about an object's relative features relies on higher-order representations (e.g., Treisman et al., 1983; see also Duncan, 1984; Kahneman et al., 1983; Vecera & Farah, 1994).

Aims of the present study

The aim of the present study was too determine whether selection of a feature-specific versus relational search mode relies on higher-order processes that determine search modes globally, or whether this is based on local adaptations. To that aim, we designed a search task for a color × size conjunction target, and examined whether it is possible to adopt a feature-specific setting for one feature of the target (e.g., size), while simultaneously tuning attention to the relative feature in another dimension (e.g., color). As a previous study already showed that conjunction search is relational (Becker et al., 2017), we enforced feature-specific search along one stimulus dimension (size or color) by randomly varying the nontargets across two displays so that the target was either not among the largest or smallest items across both displays, or not among the bluest or greenest items. We then tested whether search was relational or feature-specific in each dimension of the conjunction target by briefly presenting different irrelevant *conjunction cues* prior to the search display, and assessing which of the cues would capture attention (Fig. 1). The critical question was whether enforcing feature-specific search along one dimension of the target would lead to feature-specific search along the other target dimension as well.

If search strategies are adapted through local feedback mechanisms that adjust top-down settings independently

a Example of a Trial

b Stimuli used in Exp. 1–4
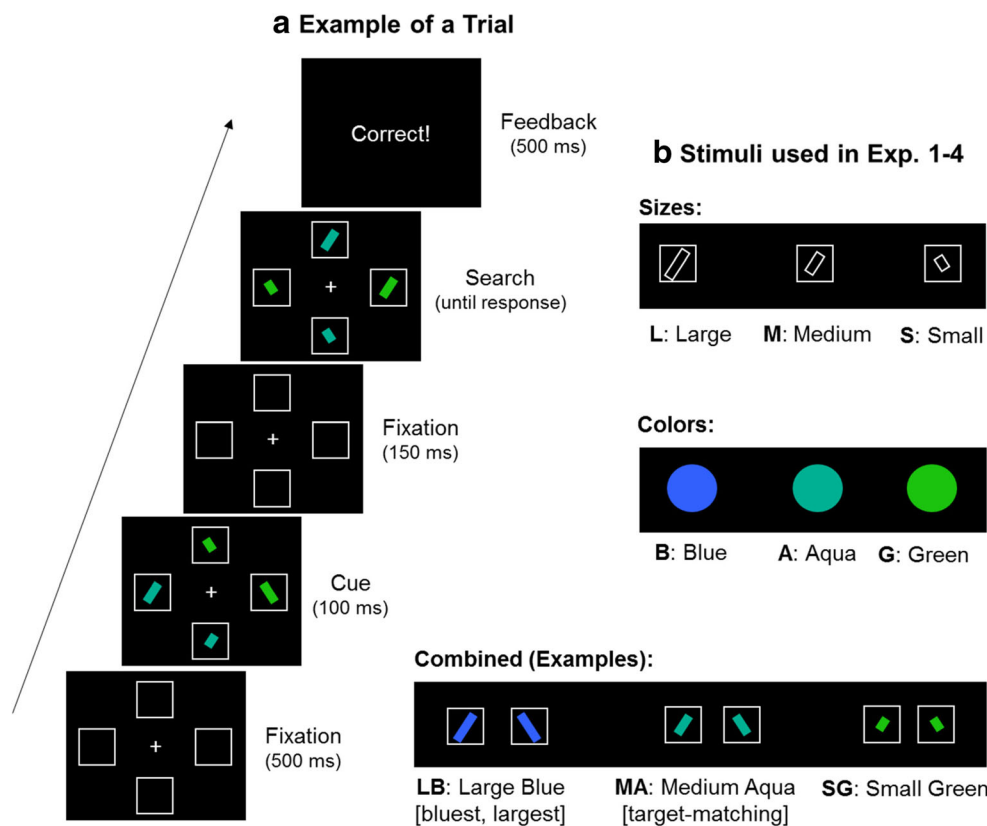
Fig. 1 a Example of a trial: Observers had to search for a medium, aqua target bar and report its orientation, while ignoring the cues. **b** Overview of stimuli used in Experiments 1–4. Search and cue displays consisted of four individually different items that had the same color or size as other items in the display (feature conjunctions). (Color figure online)

within each stimulus dimension, then tuning attention to the specific target feature should not affect how attention is tuned in the other dimension (e.g., Bundesen, 1990; Eimer & Grubert, 2014). That is, it should be possible to tune attention to the relative color whilst also tuning to the specific size of a conjunction target and *vice versa*.

On the other hand, search strategies could also be adjusted more globally to match a particular search mode across different stimulus dimensions. In this instance, encouraging feature-specific search along one dimension (e.g., color) of a conjunction target should also lead to feature-specific search along the other dimension of the target (e.g., size). Such an outcome would be expected if the decision about the search mode is made at a higher level of processing.

A third possible outcome of the present experiments is that we will find no evidence for feature-specific search in any of the conditions. Of note, attentional capture by conjunction cues has so far only been shown in one study, which allowed relational search, and in which attention was tuned to the relative features of the target (e.g., Becker et al., 2017). Moreover, in studies that enforced feature-specific tuning to the target, search for a feature conjunction was markedly less efficient than in conditions that allowed relational search (e.g., Arguin & Saumier, 2000; Brand et al., 2014; D'Zmura, 1991;

Hodsoll & Humphreys, 2001). As capture by a conjunction cue probably requires efficient search, it is possible that the conjunction cues will not attract attention when feature-specific search is enforced, and that the results will only show a marked increase in response times (RTs) or errors that reflect a serial, effortful search (e.g., Wolfe, 1994).

Overview of experiments

We probed the ability to simultaneously tune attention in two different ways (feature specific vs. relational) to a conjunction target by using a variant of the spatial cueing paradigm. Observers were instructed to search for a predefined conjunction target (e.g., medium aqua bar) and to report with a button-press response whether the target was tilted to the right or left. With this, the search-relevant feature (color × size conjunction) was independent of the response-relevant feature (orientation).

Across all experiments, the target was always a medium-sized, aqua (i.e., bluish green) item, and embedded among three nontarget items, one of which had the same color as the target but a different size (e.g., small aqua), one that had the same size but a different color (e.g., medium green), and

one that was combined of the other two nontarget features (e.g., small green). With this, the target did not have a higher feature contrast than any of the nontargets, so that bottom-up feature contrasts could not aid search (cf. Wolfe, 1994).

In Experiment 1, we presented the search display described above, in which the medium, aqua target was one of the largest, bluest items (as the nontargets were either aqua or green, and either medium or small). Correspondingly, we expected that attention would be tuned to the relative color and size of the target in Experiment 1. In Experiments 2–4, we added a second target display to the experiment that featured the same, medium aqua target in a different nontarget context, and randomly switched between displays to prevent relational search and force observers to adopt a feature-specific search setting. Specifically, in Experiment 2, the target in the second display was embedded among large or medium, aqua or green nontargets, so that the target was not consistently the largest or smallest item, but still the bluest item (across the two target displays). Together, the two target displays should enforce feature-specific tuning to the particular size of the target (medium), while still allowing relational search for the bluest item. In Experiment 3, we included medium or small, blue or aqua nontarget items in the second display, so that the target was not consistently the bluest or greenest item, but always among the largest items (in both displays). This should enforce feature-specific tuning to the particular target color (aqua) while still allowing relational search for the largest item. In Experiment 4, we enforced feature-specific tuning along both target dimensions by including large and blue nontargets in the second target display. With this, the target was neither the bluest, greenest, largest, or smallest item across both target displays, which should lead to feature-specific tuning in the color and size dimension.

To assess whether observers had indeed adopted the hypothesized search settings, we briefly presented a cue display prior to the target display that was similar to the target display, in that each cue display element was a genuine feature conjunction. Each cue display contained stimuli of two different sizes (e.g., medium, small or medium, large) and of two different colors (e.g., aqua, green or aqua, blue), and each cue was a unique combination of these features (e.g., medium, aqua; medium, blue; large, aqua; large, blue). Altogether, we presented three different cue displays in which we systematically varied the possible colors and sizes of the cue display elements, so that they could either match or mismatch the relative or feature-specific properties of the target. In Cue Displays 1 and 2, different search settings (feature specific vs. relational) should lead to attentional capture by different cues within the display.

Attentional capture was assessed by comparing the mean RT and errors to the target when each of the different cues was presented at the target location. If one of the conjunction cues in the display attracts attention, mean RT to the target should be faster (and errors should be lower) when this cue is presented at the target location than when one of the other three cues is presented at the target location (see Becker, Folk & Remington, 2010, for a similar procedure). Moreover, if attention is tuned to the specific feature values of the target (e.g., medium aqua), only cues matching the target features should attract attention, across all cue displays (i.e., only medium aqua cues). By contrast, if attention is tuned to the relative feature of the target (e.g., largest/bluest), then the cue that is among the largest/bluest elements *in the cue display* should attract attention (e.g., large, aqua cue in Cue Display 1, large, blue cue in Cue Display 2).

We also included a third, control cue display, composed of cue elements that were identical to the target and nontarget items. As the cue in this display matched both the relative color and size of the target (i.e., it was one of the largest/ bluest cues) and its physical color and size, it should attract attention both in a relational and feature-specific search mode. This cue display served as a control, against which the capture effects of the other, target-dissimilar cues could be compared.

Experiment 1

The first experiment was designed to test whether attention is indeed usually tuned to two relative features of a conjunction target. In Experiment 1, the conjunction target was a medium-sized aqua bar that was presented among three nontargets; a medium green, a small aqua, and a small green bar that together constituted the context for the target. The target color (aqua) was a mixture of green and blue, and in Experiment 1 the target was only presented among nontargets that either had the same color or were green. Similarly, the nontargets all either had the same size as the target or were smaller. Hence, the target was reliably among the bluest and largest items and could be equally well found by biasing attention to the targets' relative features ("bluest largest") or its exact feature values ("medium aqua").

To assess whether attention was biased to the exact target features or its relative features, we briefly presented one out of three possible cue displays prior to the target display (see Fig. 2). The cue displays contained four differently colored and sized cues that were unique combinations of two different colors and sizes. Each cue display contained a target-similar (aqua, medium) cue that perfectly matched the target features, and three target-dissimilar cues. In the control cue display, the three target-dissimilar cues matched the nontarget features, and thus the target-similar cue was also the bluest/largest cue in the display and would be predicted to capture both in relational and in feature-specific search. In Cue Displays 1 and 2, the target-dissimilar cues could match the relative features of the target or one relative and one specific feature of the target, and thus could distinguish between relational and feature-
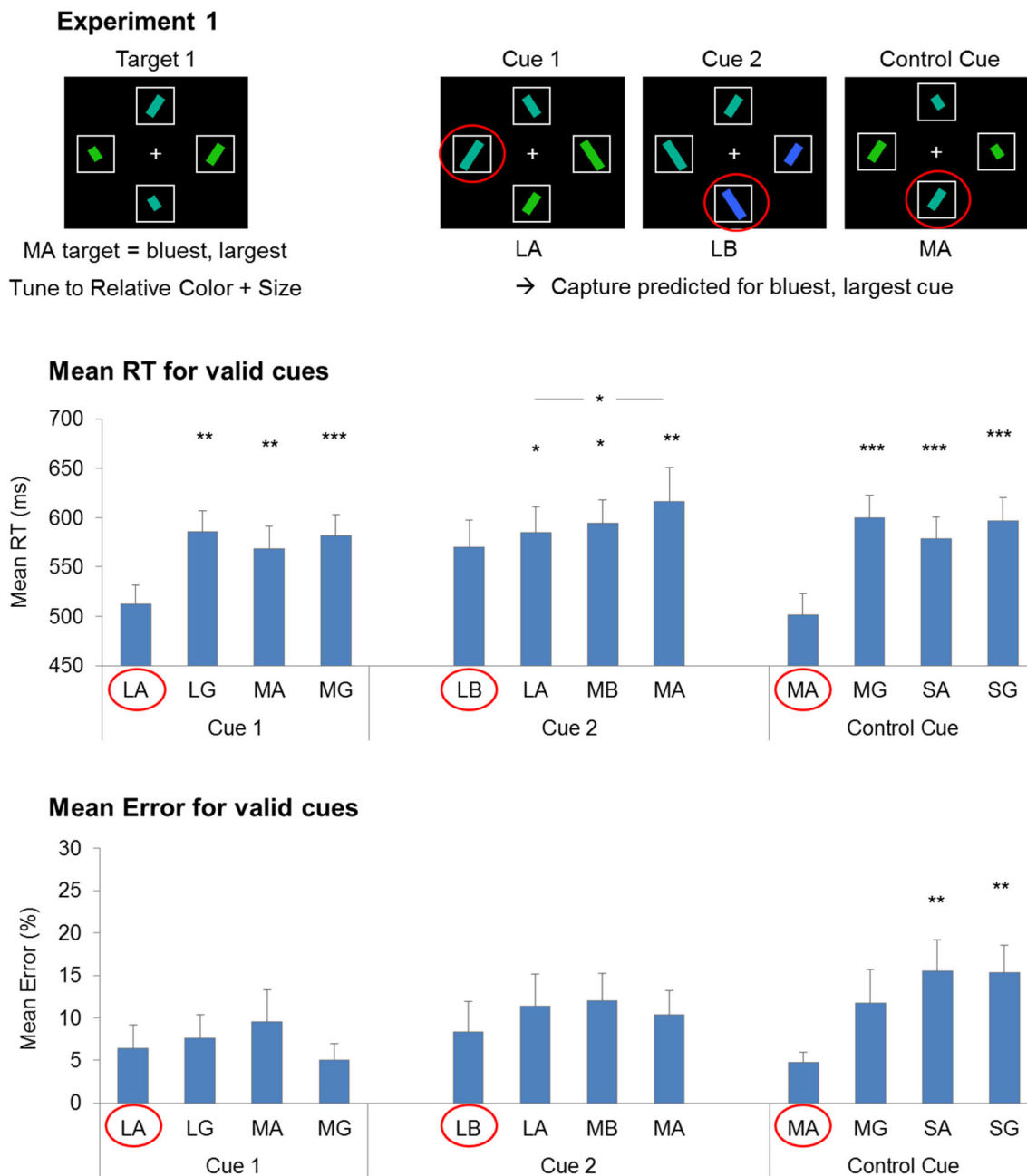
Fig. 2 The top panel shows the target display and cue displays for Experiment 1. As the MA (medium, aqua) target was consistently one of the bluest, largest item, the relational account predicts capture for those conjunction cues that are among the bluest, largest items in the respective cue displays, regardless of their similarity to the target (outlined in red). The two bottom panels depicting the results of the mean RT and errors show that these cues led to the significantly faster RT and (in part) fewer errors, as indicated by the asterisks above the other cues. Error bars represent 1 *SEM*. *$p \leq .05$. **$p \leq .01$. *** $p \leq .001$, for the comparison with the predicted cue (marked by the red circle). (Color figure online)

specific search. As shown in Fig. 2, Cue Display 1 contained a large, aqua cue in addition to the target-similar medium, aqua cue, and Cue Display 2 contained a large, blue cue in addition to the target-similar cue. Thus, if attention is tuned to the relative color and size of the target (bluest, largest), then the large cues should attract attention (i.e., large, aqua cue in the Cue Display 1, and large, blue cue in Cue Display 2), as these cues are the largest/bluest items in the respective cue displays.

However, if attention is tuned to the exact target feature values, then attention should always be attracted to the target-similar, medium aqua blue in these cue displays.

As all cue displays contained a target-similar cue, the cues that were predicted to capture in relational search also happened to be larger and bluer than the target (except in the third cue display, which served as a control). A previous study has already shown that this is irrelevant for relational search, as

capture by a cue depends only on whether it is the bluest/largest in the respective cue display, independently of whether it is larger or smaller than the target, or bluer or greener (Becker, 2010; Becker et al., 2013, 2017; Harris et al., 2013). We did not include a corresponding control condition here, as we have previously demonstrated that conjunction search is relational (i.e., independent of the physical similarities between cue and target; Becker et al., 2017), and in this study our primary objective was to distinguish between relational and feature-specific search—which required including a target-similar cue in each cue display.

Previous studies have shown a strong preference for relational search over feature-specific search in the spatial cueing task (e.g., Becker et al., 2017). However, this preference may have been (in part) because the exact feature values of the cues are difficult to judge when they appear in a different location as the search items (e.g., because the size of the target and nontargets are judged relative to the borders of the placeholder box). In the present study, we presented the cues in the same locations as the search items, to facilitate perception of the physical match vs. mismatch and remove the possible bias for relational search.

A result pattern showing relational search, with capture by the largest/bluest items in Experiment 1, could be attributed to a search asymmetry for blue or large items, in line with the observation that larger items can be found faster than smaller items (e.g., Treisman & Gelade, 1980). However, previous studies have shown that attention can be biased against such existing search asymmetries to select the smallest item in relational search (e.g., Becker, 2010; Harris et al., 2013), and a previous conjunction search study showed the same relational results pattern when the colors of the target and nontargets were flipped (requiring search for a greener target), or when color was replaced with gray values of varying luminance (darker/lighter; Becker et al., 2017). We did not include corresponding control conditions here, as searching for a smaller item may render a conjunction search task too inefficient to observe attentional capture by conjunction cues, and previous studies already rule out that relational effects are specific to certain features or stimulus dimensions.

Capture by a cue was inferred when the mean RT and errors for this cue were lower than for all other cues when each of them was presented at the target location (i.e., on valid trials; see Becker et al., 2010, for a similar procedure). We could not compute the standard validity effect for each cue, as this would have involved reusing data across different statistical tests. For example, the invalid trials of the medium aqua cue in Cue Display 1 were composed of the valid trials of the large, aqua; large, green; and medium, green cue. Hence, computing the validity effects for each cue would have involved reusing the data from each cue on four different occasions (once, to compute the valid trial performance for that cue, and then to compute the invalid trial performance for the other three cues).

Our current strategy of comparing only the valid trials for each cue ensures that the data are independent of each other and is equivalent to comparing the valid trial performance of one cue to each different trial type where this cue was invalid (see Becker et al., 2010).

Importantly, all cue displays contained a target-matching medium aqua (MA) cue. Thus, if attention is tuned to the specific color and size of the target (medium and aqua), the target-matching medium aqua cue should always show the fastest RT and/or fewest errors. By contrast, if attention is tuned to the relative size and color of the target (largest and bluest), then the cue that is among the largest/bluest cues within a given cue display should attract attention and yield the fastest RT and/or fewest errors: As shown in Fig. 2, this would be the large aqua cue in Cue Display 1, the large, blue cue in Cue Display 2, and the medium, aqua cue in Cue Display 3 (marked by a red circle in Fig. 2).

Method

Participants In each of the experiments (Experiments 1–4), we tested 13 different participants from the University of Queensland, Australia. The estimated sample size of $N = 13$ was based on a previous spatial cueing experiment that established that conjunction search is relational and used very similar stimuli and procedures as the present experiments (Becker et al., 2017; Experiments 1 and 2). Of the 13 volunteers who took part in this experiment, 10 were females, and they had a mean age of 21.6 years ($SD = 1.62$). All participants reported normal or corrected-to-normal vision, and were paid $10 for their participation.

Materials An Intel Core i7-4790 Processor 3.6 GHz computer attached to a 17-in. BenQ CRT color monitor with a resolution of $1,280 \times 1,024$ pixels and a refresh rate of 67 Hz was used to generate, present and control the stimuli displays and record responses. Subjects were seated in a normally lit room and viewed the screen from a distance of ~50 cm. Trial scheduling and response time (RT) measurement were controlled by PsychoPy software (Python Systems; Peirce et al., 2019), and all manual responses were recorded with a standard USB keyboard. A ColorCAL MKII Colorimeter (Cambridge Research Systems) was used to measure the luminance of the stimuli, and to render the stimuli equiluminant.

Stimuli The spatial cueing task consisted of four stimulus displays that were presented against a consistently black background: a fixation display, cue display, target display, and feedback display. The fixation display consisted of a central white fixation cross (1.97cm × 1.34cm; RGB: 255, 255, 255), surrounded by four square placeholder boxes outlined in white (3.8 cm × 3.8 cm; 2 pt width). The placeholders were placed at the 12 o'clock, 3 o'clock, 6 o'clock, and 9 o'clock positions,

on the outlines of an imaginary ellipse that measured 11.2 cm along the horizontal axis and 9.2 cm along the vertical axis (see Maljkovic & Nakayama, 1996, for a similar procedure).[1]

The cue display consisted of the fixation display, with the addition of four bars that were rotated 33° to the left or right from vertical and placed in the center of each box. Each bar could have one out of three different sizes: small (1 cm × 0.44 cm), medium (1.8 cm × 0.44 cm), or large (3.2 cm × 0.44 cm), and one out of three different colors: blue (RGB: 0, 150, 180; Lxy: 44.76 cd/m^2; .206, .256), aqua (RGB: 0, 153, 153; Lxy: 44.73 cd/m^2; .220, .306), or green (RGB: 0, 155, 130; Lxy: 44.72 cd/m^2; .232, .350). The cues used for the analysis of validity effects across the three cue displays in Experiment 1 were (1) a large, aqua (LA) cue presented among a large, green (LG); medium, aqua (MA); and medium, green (MG) cue; (2) a large, blue (LB) cue presented among a large, aqua (LA); medium, blue (MB); and medium, aqua (MA) cue; and (3) a medium, aqua (MA) cue presented among a medium, green (MG); small, aqua (SA); and small, green (SG) cue.

The target display was identical to the third cue display in layout, colors, and sizes of bars and placeholders. In Experiment 1, the target display always contained a medium aqua target bar that was presented together with a medium green, small aqua, and small green nontarget bar. The feedback display consisted of the written words "Correct!" or "Oops! That was wrong," which were presented in white font (Arial Black, 15 point) in the center of the screen.

Design The target position, cue display (1, 2, 3), and cue position within each display were varied pseudorandomly across trials, such that the target appeared an equal number of times with each cue in each position. This ensured that the position of the cue was uncorrelated with the position of the target, so that observers had no incentive to actively attend to the cue (25% valid trials for each of the cues). Complete counterbalancing of the positions and orientations of all uniquely colored and sized items in the cue and target display was not feasible, as this would have resulted in too many combinations. For Experiment 1, 144 cue and target displays were randomly selected such that each display contained an equal number of right-tilted and left-tilted bars, and that the critical cue and target were always presented an equal number of times at all locations. Each participant completed four runs of the 144 trials in a random order, resulting in 576 observations per participant.

Procedure Prior to the experiment, observers were instructed to search for the medium aqua bar in the target display and to

report its orientation (tilted left vs. right from vertical) using the left and right arrow keys on the keyboard. Observers were also instructed to ignore the cues that were presented prior to the target display and to respond to the target as quickly and accurately as possible, while keeping the gaze fixated on the central white fixation cross during the entire trial. To ensure that the participants had understood the task, they were asked to complete 10 practice trials that were randomly selected from the total pool of trials prior to the experiment (not recorded).

As shown in Fig 1a, each trial of the experiment started with the presentation of the fixation display for 500 ms followed by the cue display (100 ms) and the fixation display (150 ms). The target display was presented until subjects had entered their response. Immediately afterwards, the feedback display was presented. After an intertrial interval of 500 ms, in which a blank black screen was presented, the next trial started again with the presentation of the fixation display. After 288 trials, the initial instruction display was presented again, and participants were encouraged to take a short break to avoid fatigue. On average, it took 30 mins to complete the experiment.

Results

Data Anticipatory responses (RT <200 ms) or delayed responses (RT >1,200 ms) were excluded from all analyses, which led to a loss of 2.26% of all data. Unless stated otherwise, data were analyzed with repeated-measures ANOVAs, whereby we report the Greenhouse-corrected p values where necessary, together with the uncorrected degrees of freedom (for better readability). The central hypotheses were tested with two-tailed t tests. No correction for repeated testing was used, as inflated chances of finding significant differences between the conditions would not work in favor of any of the involved hypotheses, but slightly against them (because both accounts predicted that a single cue would differ from all the other cues, which in turn should *not* differ among themselves, in any given cue display).

Mean RT Figure 2 depicts the mean RT for each of the four cues in the three cue displays, when the cue was located at the target location. A 3 × 4 repeated-measures ANOVA with the variables cue display (1–3) and cue features (the four cues in each cue display, ordered from bluest/largest to smallest/greenest) showed a significant main effect of the cue display, $F(2, 24) = 4.68$, $p = .023$, $\eta_p^2 = .281$, the cue features, $F(3, 36) = 25.59$, $p < .001$, $\eta_p^2 = .681$, and a significant interaction between the two variables, $F(6, 72) = 5.26$, $p = .002$, $\eta_p^2 = .305$.

As shown in Fig. 2, the relatively matching cues that were among the largest/bluest items in the cue displays yielded the fastest responses. In Cue Display 1, responses were

significantly faster when the large, aqua (LA) cue had been presented at the target location than any of the other cues, including the large, green (LG) cue, $t(12) = 4.15$, $p = .001$, the target-matching medium, aqua (MA) cue, $t(12) = 3.27$, $p = .007$, and the medium, green (MG) cue, $t(12) = 5.02$, $p < .001$. The latter three cues did not differ significantly from each other, all $ts < 1.45$, $ps > .17$.

Similarly, in Cue Display 2, the relatively matching LB cue produced the shortest RTs, which were significantly shorter than for the large, aqua (LA) cue, $t(12) = 2.35$, $p = .035$, the medium, blue (MB) cue, $t(12) = 2.31$, $p = .039$, and the target-matching (MA) cue, $t(12) = 3.61$, $p = .004$. The latter three cues again did not differ significantly from each other, with the exception of the LA cue, which produced significantly shorter RT than the target-matching MA cue, $t(12) = 2.24$, $p = .045$.

The control cue display showed the shortest RT for the medium aqua (MA) cue, which matched both the relative and the specific color and size of the target. RTs were significantly shorter for the MA cue than for the medium green (MG) cue, $t(12) = 7.03$, $p < .001$, the small aqua (SA) cue, $t(12) = 7.82$, $p < .001$, and the small green (SG) cue, $t(12) = 6.17$, $p < .001$. By contrast, the latter three cues (MG, SA, SG) did not differ significantly from each other, all $ts < 1.69$, $ps > .11$.

Mean errors The same 3×4 ANOVA computed over the mean error scores yielded a main effect of the cue display, $F(2, 24) = 7.61$, $p = .003$, $\eta_p^2 = .388$, the cue features, $F(3, 36) = 4.33$, $p = .025$, $\eta_p^2 = .265$, and a significant interaction between the two variables, $F(6, 72) = 3.13$, $p = .032$, $\eta_p^2 = .207$.

As shown in Fig. 2 (bottom panel), the mean errors showed similar results as the mean RT. In Cue Displays 1 and 2, none of the differences in errors were significant, all $ts < 2.10$, $p \geq .058$. In the control cue display, the target-matching (MA) cue led to fewer errors than the small aqua SA cue, $t(12) = 3.53$, $p = .004$, and the small green (SG) cue, $t(12) = 3.41$, $p = .005$, whereas the difference to the medium green (MG) cue just failed to reach significance, $t(12) = 2.13$, $p = .054$. The MG, SA, and SG cues did not differ significantly from each other, $ts < 1.84$, $ps < .091$. As the observed trends and effects in the mean error scores were in the same direction as the mean RT, the results are not complicated by a speed–accuracy trade-off.

Discussion

Experiment 1 showed that the conjunction cues that matched the relative color and size of the target (bluest/largest) attracted attention most strongly, more strongly than the cues that matched the exact features of the target (medium, aqua). Capture by a relatively matching conjunction cue was observed even when it did not share any of the target features (e.g., large, blue cue; in Cue Display 2), and when it directly

competed against another cue that had the same features as the target (medium, aqua; see Cue Displays 1 and 2). These results are inconsistent with a feature-specific search mode, and show that attention was tuned to the relative features of the target (bluest, largest), in line with earlier results (Becker, 2017).

Of note, the large, blue (LB) cue seemed to attract attention less strongly than the target-matching (MA) cue in Cue Display 3, or the large, aqua (LA) cue in Cue Display 1, as RTs were slower for the LB cue. Two-tailed t tests confirmed that the LA and MA cues in Cue Displays 1 and 3 speeded responses more than the LB cue in Cue Display 2, $ts > 3.18$, $ps \leq .008$, whereas the LA cue did not differ from the MA cue, $p = .54$. At a first glance, it may be tempting to think that these results may reflect feature-specific tuning to the exact color of the target. However, this seems unlikely, as Cue Display 2 also contained a target-matching MA cue and a large, aqua (LA) cue; yet the LB cue produced significantly shorter RTs than these other two cues. Moreover, in Cue Display 2, the target-matching MA cue produced the longest RT, which rules out that the differences were due to a feature-specific setting for the exact size or color of the target. The results also rule out optimal tuning for a color between blue and aqua: If attention had been tuned to a shifted target color, the target-similar MA cue and the other aqua-colored cues should still have shown an advantage over the other cues. Instead, the MA cue showed no advantage to the nontarget-colored, green cues in Cue Display 1, and had the longest RT in Cue Display 2.

Instead, stronger capture by the MA and LA cues than by the LB cue may be because the blue color was too similar to the aqua color, so that the large, blue cue was maybe not always immediately identified as the bluest cue. While this explanation is still speculative, the results showed that conjunction cues with very different features captured attention across the different displays (LA, LB, and MA), which demonstrates that capture did not depend on the similarity of the cue to the target features, contrary to current feature similarity accounts (e.g., Duncan & Humphreys, 1989; Martinez-Trujillo & Treue, 2004; Treisman & Sato, 1990). Similarly, the observation that a cue with target-dissimilar features attracts attention is inconsistent with guidance by a holistic target representation (e.g., Duncan, 1984; Kahneman et al., 1983; Treisman et al., 1983; Vecera & Farah, 1994), and instead support the relational view that attention is guided by the relative features of the target (e.g., Becker, 2010).

With this, the results replicate and extend the results of a previous study (Becker et al., 2017), by showing (1) that observers preferentially tune attention to the relative features of a conjunction target, even if the conditions allow localizing the target equally well in terms of its physical or relative features, and (2) by showing that an irrelevant conjunction cue can attract attention in the spatial cueing paradigm, despite the fact that the cue does not have a unique feature or any relevant

feature contrast and thus is not salient (e.g., Becker et al., 2017).

However, the two main research questions are still open, viz., whether tuning to the specific feature(s) of a conjunction target will still allow efficient localization of the target, and whether it is possible to simultaneously adopt two different search modes (feature specific vs. relational) in search for the conjunction target.

Experiment 2

The aim of Experiment 2 was to test whether attention can be simultaneously biased to the specific feature value of the target in one dimension, and its relative feature in the other dimension. Specifically we tested whether observers can simultaneously adopt a feature-specific setting for the size of the conjunction target (i.e., tune to medium), and a relational set for the color (i.e., tuning to bluer or the bluest item).

In Experiment 2, a feature-specific search strategy for the target size was encouraged by using two target displays. In addition to the target display used in Experiment 1, in which the target was among the bluest, largest item, we used a second target display, in which the target was the same (medium, aqua), and presented among a large, aqua; medium, green; and a large, green nontarget. Thus, across the two target displays, the target was still consistently among the bluest items, but its relative size varied between being larger (than the small nontargets: Target Display 1) or smaller (among the large nontargets: Target Display 2).

Keeping the physical size of the target constant while randomly varying its relative size renders relational search impossible and enforces a feature-specific search strategy, in which attention is biased to the physical target size (e.g., Becker, Harris, Venini, & Retell, 2014b; Harris et al., 2013). A hallmark of feature-specific tuning is that only target-similar cues attract attention (e.g., Harris et al., 2013). We used the same cue displays as in Experiment 1. The critical question was whether we can simultaneously adopt a feature-specific search for the size of the conjunction target while simultaneously conducting a relational search for the color of the conjunction target (bluest).

If attention can be tuned to the specific size of the target and to the relative color, then in Experiment 2 the cue with a medium size that is among the bluest cues in the cue display should attract attention. Thus, in Cue Display 1, now the *medium, aqua* (and not the large, aqua) cue should attract attention, as the medium, aqua cue has a medium size and is among the bluest items; and in Cue Display 2, now the *medium, blue* cue should attract attention (not the large, blue cue; marked by a red circle in Fig. 3). In the control-cue display, the medium, aqua control cue should still attract attention, as it matched both the relative features of the target and its physical features.

A corresponding result would indicate that attention was relationally tuned to the bluest item, while simultaneously being tuned to the specific target size (medium).

Alternatively, it is conceivable that attention cannot be tuned in two different ways (feature specific, relational) to different features of a conjunction target. It is possible that once we enforce tuning to the specific size of the target, attention will also be tuned to the specific color of the conjunction target. Such an outcome would be predicted if the search mode or search strategy is determined by higher-order processes that adjust search strategies more globally, across different stimulus dimensions. In this case, only the target-matching, medium aqua cues should attract attention across all cue displays (Cue Displays 1–3), as they matched the physical color and size of the target.

As in Experiment 1, attentional capture by the cues was assessed by comparing the mean RT and errors when each different cue was presented at the target location.

Of note, it is possible that the results are not due to top-down tuning of attention to the target, but automatic priming effects that bias selection to the features of the target selected on the previous trial (or its relative features; e.g., Becker, 2007, 2008a, 2008b, 2010; Folk & Remington, 2008; Leonard & Egeth, 2008; Maljkovic & Nakayama, 1994). As Experiment 2 included two target displays, this possibility could now be assessed by comparing the data on repeat trials (in which the target display on the previous trial was the same) versus switch trials (in which the target display on the previous trial was different). If selection of a particular conjunction cue (or selection of the conjunction target) is due to intertrial priming effects, capture by the conjunction cue should only occur on repeat trials, not on switch trials.[2]

Method

Participants Thirteen new, naïve volunteers from the University of Queensland, Australia, took part in this experiment (10 females, mean age = 26.1 years, *SD* = 7.90).

Materials The materials were the same as those used in Experiment 1.

Stimuli The stimuli were identical to those of Experiment 1, with the following exception: In addition to the target display used in Experiment 1 (which included a medium, aqua target among a medium, green; a small, aqua; and a small, green nontarget), a second target display was used that consisted

---

[2] We also analyzed the data separately for trials in which a given cue display was preceded by Target Display 1 versus Target Display 2, across all experiments. The results were very similar, indicating that capture by a given cue was not mediated by the preceding target display, but depended on a more durable search setting that had been adopted to allow selection across both target displays. (Data available on request.)
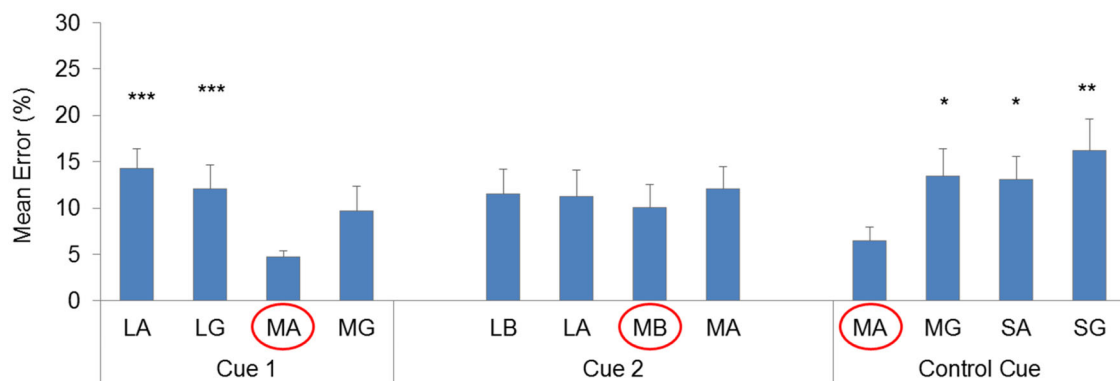
Experiment 2
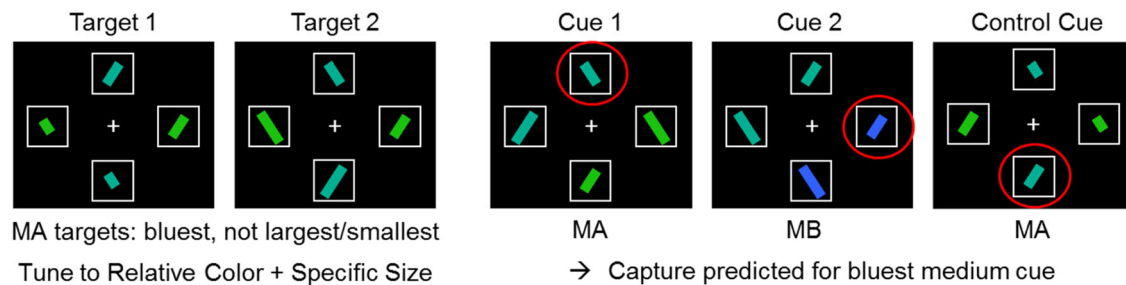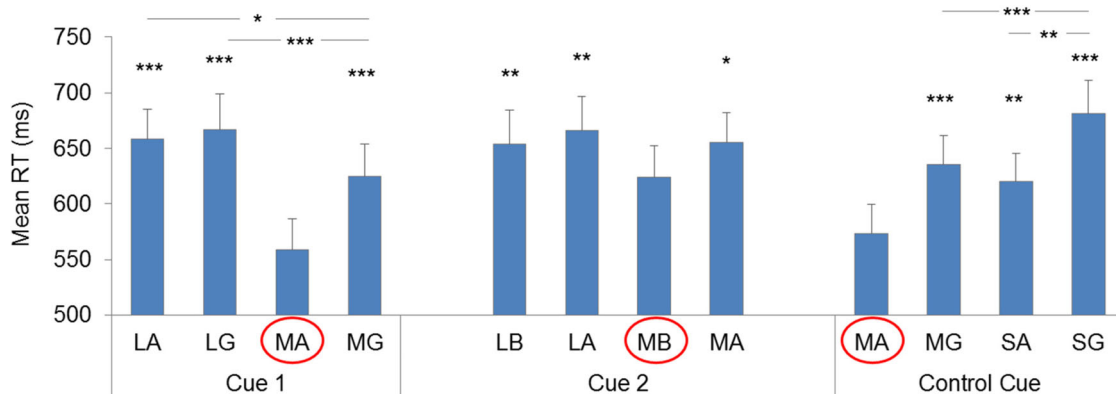
Fig. 3 Predictions for Experiment 2 (top panel), mean RT (middle panel), and error scores (bottom panel) for each cue when it was presented at the target location. As predicted, the cues matching the exact size of the target (medium) and its relative color (bluer) reliably produced the shortest RT and lowest errors, reflecting that they attracted attention. Error bars represent 1 *SEM*. *$p \leq .05$. **$p \leq .01$. ***$p \leq .001$. Asterisks on top of the histograms are for the comparison with the predicted cue (marked by the red circle). (Color figure online)

of a medium, aqua target, presented together with a large, aqua; a medium, green; and a large, green nontarget.

Design The design of Experiment 2 was identical to Experiment 1, with the addition of the second target display. Each of the 144 cue displays used in Experiment 1 were presented prior to each of the two target displays using the same methods for counterbalancing the target and cue position, cue displays, and tilt direction of the bars, resulting in 288 trials.

Each participant completed two runs of these in random order, resulting in 576 trials per participant.

Procedure The procedure for Experiment 2 mirrored that of Experiment 1. As with Experiment 1, participants were instructed to search for a medium, aqua bar and respond to its orientation by using the left and right arrow keys on the computer keyboard. No instructions were issued regarding specific search strategies: The target features were specified

in the same way as in Experiment 1 ("Search for the medium aqua bar."). As in Experiment 1, subjects completed 10 practice trials prior to the experiment that were not recorded, to ensure that they had understood the task.

Results

Data Applying the same exclusion criteria as in Experiment 1 led to a loss of 3.37% of all data in Experiment 2.

Mean RT The RT data were first analyzed by a $3 \times 4 \times 2$ repeated-measures ANOVA comprising the variables cue display (1–3), cue features (the four cues in each display ordered from bluest, largest to smallest, greenest), and priming (target display repeated from previous trial vs. switched). The results showed a significant main effect of cue display, $F(2, 24) = 13.11$, $p = .001$, $\eta_p^2 = .522$, of the cue features, $F(3, 36) = 21.54$, $p < .001$, $\eta_p^2 = .642$, and a significant interaction between these variables, $F(6, 72) = 18.78$, $p < .001$, $\eta_p^2 = .610$. There was no main effect of priming, and priming did not interact with any of the variables, all Fs < 1.69, all ps $> .12$. Hence, for the subsequent analyses, the data were pooled over repeat and switch trials.

As shown in Fig. 3, the data showed feature-specific tuning to the size of the target (medium), and relational tuning to the color (bluest), as the bluest cue that matched the specific size of the target produced the shortest RT. For Cue Display 1, the medium, aqua (MA) cue produced significantly shorter RTs than the large, aqua (LA) cue, $t(12) = 8.11$ $p < .001$, the large, green (LG) cue, $t(12) = 8.09$ $p < .001$, and the medium, green (MG) cue, $t(12) = 6.59$, $p < .001$, indicating that it captured attention. In addition, the cues that did not capture attention differed from each other, with the MG cue producing shorter RT than the LG cue, $t(12) = 4.83$, $p < .001$, and the LA cue, $t(12) = 2.92$, $p = .013$, whereas the LG and LA cues did not differ from each other, $t < 1$.

Critically, in Cue Display 2, the medium, blue (MB) cue that matched the relative color, but the physical size of the target produced the fastest responses, which were significantly faster than those for the large, blue (LB) cue, $t(12) = 3.43$, $p = .005$, the large, aqua (LA) cue, $t(12) = 4.12$, $p = .001$, and the target-matching MA cue, $t(12) = 2.82$, $p = .015$. The latter three cues (LB, LA, MA), which failed to match either the relative color of the target (not bluest) or its physical size (not medium), did not differ from each other, all ts < 1.55, $p > .14$.

In the control cue display, the target-matching MA cue led to significantly shorter RT than the medium, green (MG) cue, $t(12) = 6.94$, $p < .001$, the small, aqua (SA) cue, $t(12) = 3.94$, $p = .002$, and the small, green (SG) cue, $t(12) = 6.61$, $p < .001$. The long RTs associated with the SG cue were also significantly longer than the RTs of both the MG cue, $t(12) = 3.68$, p

$= .003$, and the SA cue, $t(12) = 4.81$, $p < .001$, whereas the MG and SA cues did not differ from each other, $t(12) = 1.34$, $p = .207$.

Mean errors The same ANOVA computed over the mean errors showed only a marginally significant effect of the cue features, $F(3, 36) = 3.20$, $p = .057$, $\eta_p^2 = .211$, and a significant interaction of this variable with cue displays, $F(6, 72) = 5.07$, $p = .005$, $\eta_p^2 = .297$, but no significant effect of priming or interaction with priming, all Fs < 1.46, ps $> .24$. Hence, for the subsequent analyses, the error data were pooled over repeat and switch trials.

As shown in Fig. 3, the mean errors similarly showed the lowest error rates for the cue that matched the relative color of the target (bluest) and the physical size (medium): In Cue Display 1, the errors for the target-matching MA cue were significantly lower than for the LA cue, $t(12) = 4.82$, $p < .001$, and the LG cue, $t(12) = 3.35$, $p = .006$, and marginally significantly lower than for the MG cue, $t(12) = 2.03$, $p = .065$. In addition, the difference between the LA and MG cues approached significance, $t(12) = 1.92$, $p = .079$, whereas the other cues did not differ, all ts < 1.2, ps $> .28$. In Cue Display 2, the cues did not differ significantly from each other, all ts < 1.6, ps $> .15$. In the control cue condition, the target-matching MA cue was associated with the lowest error scores, significantly lower than for the MG cue, $t(12) = 2.25$, $p = .045$, the SA cue, $t(12) = 2.47$, $p = .029$, and the SG cue, $t(12) = 3.20$, $p = .008$. The latter three cues (MG, SA, and SG) did not differ from each other, all ts < 1.36, ps $> .20$. With this, the results of the mean errors were similar to the RT results, so that interpretation of the data is not complicated by a speed–accuracy trade-off.

Discussion

The results of Experiment 2 establish, for the first time, that observers can simultaneously adopt a feature-specific setting along the dimension of size and a relational setting along the dimension of color of a conjunction target. This is borne out by the results showing attentional capture by the cues that had the same size as the target (medium; feature-specific tuning), and were among the bluest cues in the display, regardless of their exact color (aqua or blue; relational tuning). Deviating from Experiment 1, which showed capture by the large, aqua (LA) cue in Cue Display 1, Experiment 2 enforced top-down tuning to the exact size of the target and correspondingly, we instead observed capture by the medium, aqua (MA) cue in Cue Display 1. Similarly, Cue Display 2 showed capture by the medium, blue (MB) cue, rather than by the large, blue cue (LB; as in Experiment 1), showing that color was still selected based on its relative feature (bluer), whereas size was selected based on its specific feature value (medium).

Method

Participants Thirteen new and naïve volunteers from the University of Queensland, Australia, took part in this experiment (10 females, mean age = 22.2 years, SD = 1.65).

Materials The materials were the same as those used in Experiment 1.

Stimuli, design, and procedure The stimuli, design, and procedure were the same as in Experiment 1 and 2, with the following exceptions. First, we had to adjust the colors for this experiment and Experiment 4, because pilot testing revealed that some participants could not find the target with above-chance accuracy when we used the same colors as in Experiments 1 and 2. Specifically, the colors (blue, aqua, green) had to be more dissimilar from each other, resulting in a higher color contrast, to allow participants to successfully detect the aqua target when it was randomly presented among green or blue nontargets. The requirement for a target with a larger color contrast is unsurprising, as it has previously been found that feature specific search for a color target is less efficient than relational search (Becker, Harris, Venini, & Retell, 2014b). To increase discriminability of the colors, the colors used in Experiments 3 and 4 were changed to the following hues: blue (RGB: 66, 193, 255; Lxy: 70.56 cd/m^2; .206, .237), aqua (RGB: 13, 195, 168; Lxy: 70.26 cd/m^2; .230, .348), and green (RGB: 0, 204, 0; Lxy: 70.34 cd/m^2; .295, .593). The size of the stimuli remained the same as in Experiments 1 and 2.

The first of the two target displays used in Experiment 3 was identical to that used in Experiment 1 (except for the adjusted colors) and showed a medium, aqua target among a medium, green;, a small, aqua; and a small, green nontarget bar (i.e., target one of the bluest, largest items). In the second display, the target was a medium, aqua bar, presented among a small, aqua; medium, blue; and small, blue nontarget bar (target one of the greenest, largest items), to prevent tuning to the bluest item. Across the two target displays, we thus enforced tuning to the exact target color (aqua), while allowing tuning to the relative size. The cue displays in Experiment 3 were the same as those used in the previous experiments (except for the adjusted colors).

Results

Data Applying the same exclusion criteria as in Experiments 1 and 2 (RT <200 ms, RT > 1,200 ms) led to a loss of 1.72% of all data.

Mean RT First, to analyze the data for possible priming effects, we computed the same $3 \times 4 \times 2$ repeated-measures ANOVA with the variables cue display (1–3), cue features (the four cues in each display ordered from bluest, largest to smallest, greenest), and priming (target display repeated from previous trial vs. switched) over the mean RT. The results showed a significant main effect for the cue display, $F(2, 24) = 9.30$, $p = .003$, $\eta_p^2 = .436$, the cue features, $F(3, 36) = 22.64$, $p < .001$, $\eta_p^2 = .654$, and a significant interaction between cue display and cue features, $F(6, 72) = 13.64$, $p < .001$, $\eta_p^2 = .532$ (all other Fs < 1.98, ps > .18). As the results were not modulated by priming, we pooled the data across trials in which the target display was repeated versus switched.

Inspection of the data (see Fig. 4) revealed that the cues that matched the relative size of the target (largest) and its specific color (aqua) led to the fastest responses in Experiment 3. In Cue Display 1, the LA cue produced the shortest RT, significantly shorter than for the LG cue, $t(12) = 6.69$, $p < .001$, the target-matching MA cue, $t(12) = 3.77$, $p = .003$, and the MG cue, $t(12) = 5.88$, $p < .001$. By contrast, the latter three cues (LG, MA, MG) did not differ from each other, all ts < 1.54, ps > .15.

In Cue Display 2, the LA cue led to significantly shorter RT than the LB cue, $t(12) = 10.29$, $p < .001$, the MB cue, $t(12) = 9.15$, $p < .001$, and the target-matching MA cue, $t(12) = 3.48$, $p = .005$. In addition, RT were significantly shorter for the target-matching MA cue than for the LB cue, $t(12) = 4.06$, $p = .002$, and the MB cue, $t(12) = 5.79$, $p < .001$, whereas the MB and LB cue did not differ from each other, $t(12) = 1.39$, $p = .19$.

In the control cue display, the target-matching MA cue led to significantly shorter RTs than the MG cue, $t(12) = 7.69$, $p < .001$, the SA cue, $t(12) = 7.83$, $p < .001$, and the SG cue, $t(12) = 7.60$, $p < .001$. In addition, RTs were shorter for the SA cue than for the SG cue, $t(12) = 2.92$, $p = .013$ (all other ts < 1.82, p > .094).

Mean errors The same $3 \times 4 \times 2$ ANOVA computed over the mean error scores revealed only a significant main effect of cue display, $F(2, 24) = 8.95$, $p = .001$, $\eta_p^2 = .427$, and a significant interaction between cue display and cue features, $F(6, 72) = 6.80$, $p = .001$, $\eta_p^2 = .362$ (all other Fs < 1.5, ps > .111).

Similar to the mean RTs, the fewest errors were committed for the aqua conjunction cue that was among the largest cues in the cue display (see Fig. 4, bottom): In Cue Display 1, there were fewer errors for the LA cue than for the LG cue, $t(12) = 2.53$, $p = .027$, and the target-matching MA cue, $t(12) = 3.12$, $p = .009$, whereas the other cues did not differ significantly from each other, all ts < 1.05, ps > .31. In Cue Display 2, the LA cue led to significantly fewer errors than the LB cue, $t(12) = 3.17$, $p = .008$, the MB cue, $t(12) = 3.46$, $p = .005$, and the MA, $t(12) = 6.12$, $p < .001$. The latter three cues (LB, MB, and MA cues) did not differ, all ts < 1.15, ps > .27. In the control cue display, the target-matching MA cue led to
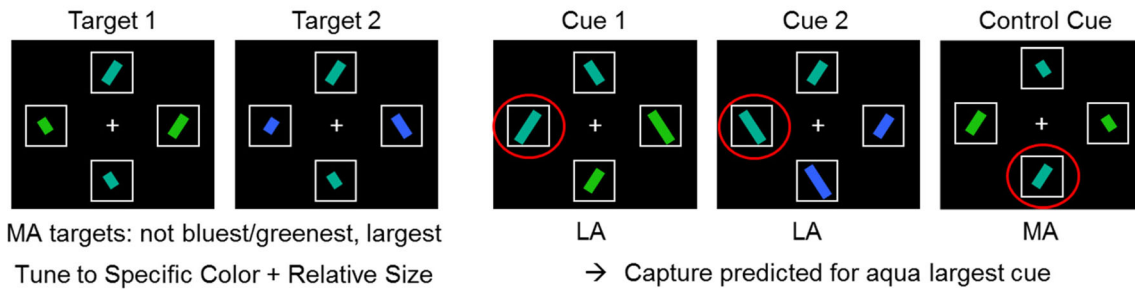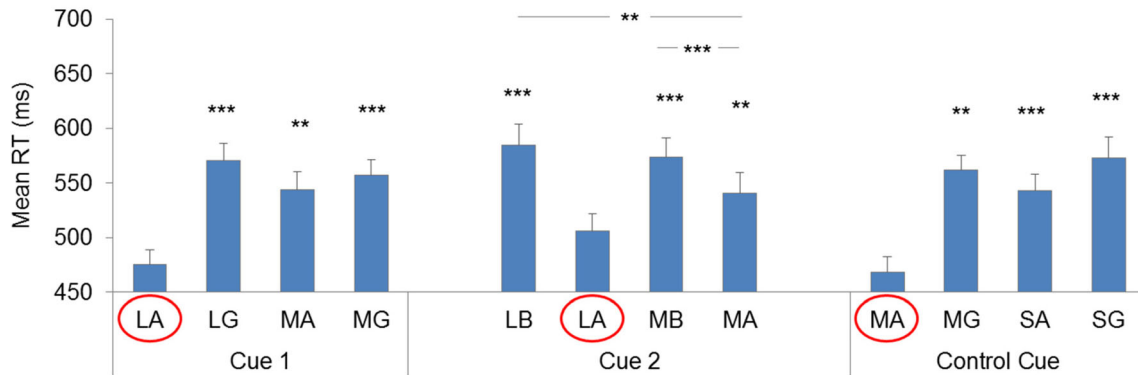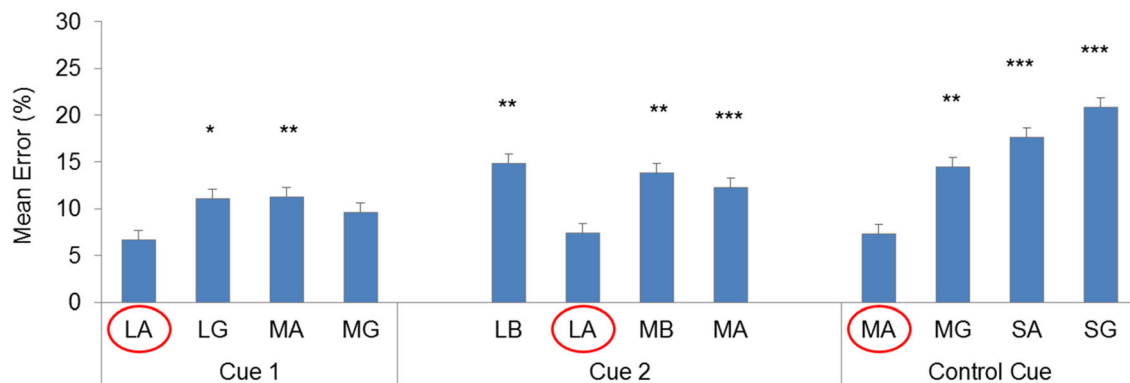
Fig. 4 Predictions and results for Experiment 3, in which we enforced a feature-specific search for the target color (aqua) while allowing relational search along the size dimension (largest). As predicted, the cues matching the relative size of the target (largest) and its specific color (aqua) now reliably attracted attention, as reflected in the fact that they produced the shortest RT and lowest errors. Error bars represent 1 *SEM*. *$p \leq .05$. **$p \leq .01$. ***$p \leq .001$. Asterisks on top of the histograms are for the comparison with the predicted cue. (Color figure online)

significantly fewer errors than the MG cue, $t(12) = 3.81$, $p = .002$, the SA cue, $t(12) = 7.62$, $p < .001$, and the SG cue, $t(12) = 4.64$, $p = .001$, whereas the latter three cues (MG, SA, SG) did not differ from each other, all $ts < 2.0$, $ps > .068$. Taken together, these results show that attention was tuned to the largest, aqua item, demonstrating that observers had adopted a relational search setting for the size of the conjunction target (largest), and a feature-specific search for its color (aqua).

Discussion

Experiment 3 showed that observers could simultaneously adopt a feature-specific search mode along the dimension of color and a relational search mode along the dimension of size of a conjunction target. This was reflected by the fact that the three cues that matched the exact color of the target (aqua) and its relative size (largest) all produced the fastest RTs and/or

lowest error scores when they were presented at the target position, indicating that they all reliably attracted attention (large, aqua; large, aqua; and medium, aqua for Cue Displays 1–3, respectively). For instance, capture by the large, aqua (LA) cue in Cue Display 2 was observed despite the fact that the cue display also contained a blue cue (large, blue cue) and a cue that perfectly matched the target (medium, aqua). The fact that (deviating from Experiments 1 and 2) the LA (large, aqua) cue now attracted attention shows that attention was tuned not to the relative size of the target (largest) but to its specific color value (aqua).

Comparing the mean RTs between the cues that attracted attention showed that the MA cue in the control display led to faster responses than the LA cue in Cue Display 2, $t(12) = 3.18$, $p = .008$, whereas the differences between the MA cue and the LA cue in Cue Display 1 just failed to reach significance, $t(12) = 2.0$, $p = .068$. Deviating from Experiments 1 and 2, the MA cue also led to the second fastest responses in Cue Displays 1 and 2 (after the LA cue; see Fig. 4), which indicates that the MA cue may have attracted attention on a portion of trials. However, the MA cue only differed significantly from the other nonmatching cues in the mean RT of Cue Display 2 (not in the errors or the RT or errors of Cue Display 1), indicating that it only weakly attracted attention (if at all). By comparison, the capture effect of the LA cues in Cue Display 1 and 2 were quite large and led to significantly shorter RT and significantly fewer errors than the target-matching MA cue in the same display. Thus, the results reflect that attention was predominantly tuned to the relative size of the target, while simultaneously being tuned to its specific color.

The fact that the colors had to be rendered more dissimilar in Experiment 3 to allow feature-specific tuning to the color of the target is in line with previous reports that feature-specific tuning renders search less efficient than tuning to the target's relative color (Becker, Harris, Venini, & Retell, 2014b). In the context of the present experiments, the requirement to adjust the colors also reinforces the view that capture by the blue cues in Experiments 1 and 2 was reduced because blue was quite similar to aqua—which renders the present results more compatible with previous findings that showed equally large capture effects for relationally matching cues, regardless of the specific color (blue, aqua) or the similarity to the target color (e.g., Becker et al., 2017).

Experiment 4

Experiment 4 was designed to test whether attention can be simultaneously tuned to the specific color and size of the target (using the colors created for Experiment 3). As mentioned before, previous studies often interpreted the results of a conjunction search task as reflecting feature-specific effects, but

did not test whether attention was biased to the relative features or the exact feature values of the target (e.g., Brand et al., 2014; Treisman & Sato, 1990; Wolfe, 1994, 1998). Given that there appears to be a preference for relational tuning, it is not clear whether attention can be simultaneously tuned to the two feature values of a target in conjunction search in in a manner that will evoke contingent capture by an irrelevant conjunction cue.

To address this question, Experiment 4 encouraged feature-specific search along both dimensions of color and size. Specifically, the target displays were constructed such that the target was always a medium, aqua item, but neither the bluest nor greenest in color, and neither the largest nor smallest in size across the two displays.

The first target display, in which the target was among the bluest/largest items, was identical to the one used in Experiment 3. The second target display showed the medium, aqua target among a medium, blue; large, aqua; and large, blue nontarget. In this second target display, the target is relatively one of the greenest, smallest items. Thus, the relative feature of the target varied randomly between larger and smaller, and greener and bluer, which rendered a relational search strategy unviable, and encouraged tuning attention to the specific feature values of the target (medium aqua).

If attention can be efficiently biased to the physical color and size of a conjunction target, then only the cues matching both the target's color and size should attract attention. Therefore, across all cue displays (1–3), now the medium, aqua cue (MA) should facilitate responding most strongly, leading to the shortest RTs and/or the fewest errors. A corresponding result would show that we can effectively tune attention to two exact feature values of a conjunction target, in a way that (only) target-similar conjunction cues can attract attention, thus providing evidence for a genuine similarity effect in conjunction search (that is not due to tuning to the relative features of the conjunction cue; Duncan & Humphreys, 1989).

Method

Participants Thirteen naïve volunteers from the University of Queensland, Australia, took part in this experiment (nine females, mean age = 20.4 years, $SD = 1.54$).

Materials The materials were the same as those used in Experiment 1.

Stimuli, design, and procedure The stimulus colors were identical to those of Experiment 3 (as Experiment 4 also required feature-based search for the specific color of the conjunction target). The two target displays used in Experiment 4 included the target display used in the previous experiments, which consisted of a medium, aqua target among a medium, green; small, green; and small, aqua nontarget (target among the

largest/bluest items). The second target display contained a medium, aqua target, presented with a large, aqua; a medium, blue; and a large, blue nontarget. In this display, the target was among the smallest/greenest items, which prevented relational search and enforced tuning to the exact feature values of the target (medium, aqua). The cue displays were identical to those used in Experiment 3. All aspects of the design and procedure were identical to those used in Experiment 3.

Results

Data Applying the same exclusion criteria as in Experiments 1–3 led to a loss of 1.87% of all data in Experiment 4.

Mean RT A 3 × 4 × 2 repeated-measures ANOVA with the variables cue display (1–3), cue features (the four cues in each display ordered from bluest, largest to smallest, greenest), and priming (target display repeated from previous trial vs. switched) computed over the mean RTs showed only a significant main effect for the cue features, $F(3, 36) = 10.53, p < .001$, $\eta_p^2 = .467$, and a significant interaction between cue display and cue features, $F(6, 72) = 19.76, p < .001, \eta_p^2 = .622$, as well as a 13-ms priming effect, $F(1, 12) = 15.73, p < .001, \eta_p^2 = .562$, but no other significant effects or interactions, all other Fs < 1, ps > .41. As priming did not interact with any of the variables, we pooled the data across repeat and switch trials.

As shown in Fig. 5, when the target displays enforced tuning of attention to the specific color and size of the target, the target-matching MA conjunction cues produced the shortest RTs. In Cue Display 1, the MA cue led to shorter RT than the LA cue, $t(12) = 6.41, p < .001$, the LG cue, $t(12) = 7.64, p < .001$, and the MG cue, $t(12) = 6.12, p < .001$. In addition, the MG cue produced significantly shorter RTs than the LG cue, $t(12) = 2.22, p = .047$ (all other ts < 1.08, ps > .30). In Cue Display 2, the target-matching MA cue produced significantly faster responses than the LB cue, $t(12) = 7.91, p < .001$, the LA cue, $t(12) = 7.02, p < .001$, and the MB cue, $t(12) = 10.57, p < .001$. In addition, responses were faster for the LA cue than for the LB cue, $t(12) = 3.90, p = .002$, and the MB cue, $t(12) = 4.09, p = .002$, whereas the LB and MB conjunction cues did not differ from each other, $t < 1$. The control cue showed similar results, with the MA cue producing the fastest responses, significantly faster than the MG cue, $t(12) = 4.08, p = .002$, the SA cue, $t(12) = 3.39, p = .005$, and the SG cue, $t(12) = 8.71, p < .001$. In addition, responses were faster for the SA cue than for the SG cue, $t(12) = 4.99, p < .001$, and for the MG cue, $t(12) = 2.47, p = .029$, whereas the MG and SA cues did not differ, $t(12) = 1.19, p = .26$.

Mean errors The same 3 × 4 × 2 ANOVA computed over the mean error scores revealed only a significant main effect for the cue displays, $F(2, 24) = 8.95, p = .001, \eta_p^2 = .427$, and a

significant interaction between cue features and cue display, $F(6, 72) = 6.80, p = .001, \eta_p^2 = .362$ (all other Fs < 1.5, ps > .111).

As shown in Fig. 5, errors for the target-matching MA cue in Cue Display 1 were significantly lower than for the LA cue, $t(12) = 4.02, p = .002$, but not significantly lower than for the LG and MG cues. However, the LA cue led to significantly more errors than the LG cue, $t(12) = 2.69, p = .020$, and the MG cue, $t(12) = 3.80, p = .003$. In Cue Display 2, error rates for the target-matching MA cue were only significantly lower than for the MB cue, $t(12) = 3.24, p = .007$, not lower than for the LB or LA cues. However, errors were lower for the LB cue than for the MB cue, $t(12) = 2.57, p = .024$ (all other ts < 1.65, ps > .12). In the control cue display, the target-matching MA cue only produced fewer errors than the SA cue, $t(12) = 2.92, p = .013$. Apart from that, the differences between the MG and SG cue were significant, $t(12) = 2.38, p = .035$ (all other ts < 1.88, ps > .084).

Discussion

Experiment 4 showed capture only by the target-matching conjunction cue, which means that attention was tuned to the specific color and the specific size of the target, not its relative features. Deviating from the previous experiments, Experiment 4 showed a small, 13-ms priming effect. Importantly, priming did not modulate capture, as this would have led to a significant interaction between priming and the different cues and/or cue displays, contrary to the findings. Hence, the results suggest that priming only speeded responses to the target on repeat trials, without modulating capture by the cues.

Interestingly, comparing the capture effects across the different cue displays showed significantly faster RTs for the target-similar MA cue in Cue Display 1 than in Cue Display 3, $t(12) = 4.95, p < .001$, whereas the corresponding differences between Cue Displays 1 and 2 just failed to reach significance, $t(12) = 2.0, p = .069$. This is an interesting finding, because Cue Displays 2 and 3 were identical to the two target displays, whereas Cue Display 1 did not match the target displays (as the context cues did not match the nontarget items). The finding that the cue in Cue Display 1 produced the strongest capture effect rules out that capture in Experiment 4 was mediated by other forms of attentional capture, such as capture by display-wide features (Gibson & Kelsey, 1998), or inhibition of those nontarget features that do not match the target features (e.g., Treisman & Sato, 1990). If capture were mediated by a match in display-wide features or nontarget inhibition, the cue displays that were identical to (one of) the target displays should have elicited stronger capture, contrary to observations. Thus, it seems safe to conclude that selection of the target-similar MA cue was due to feature-specific tuning to the color and size of the conjunction target.
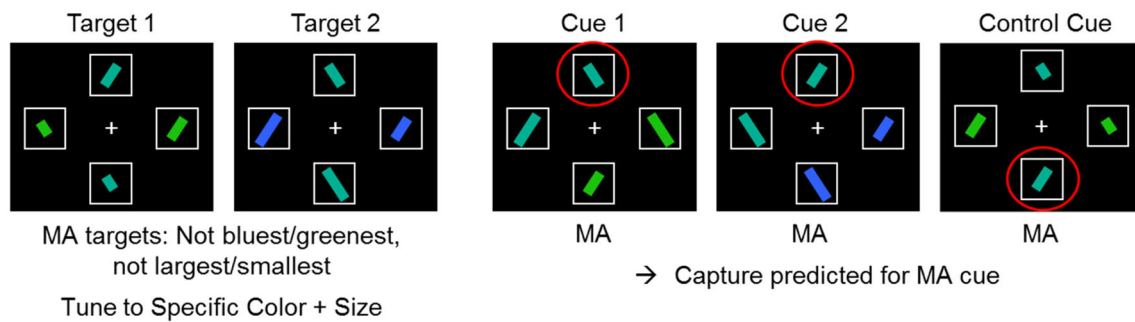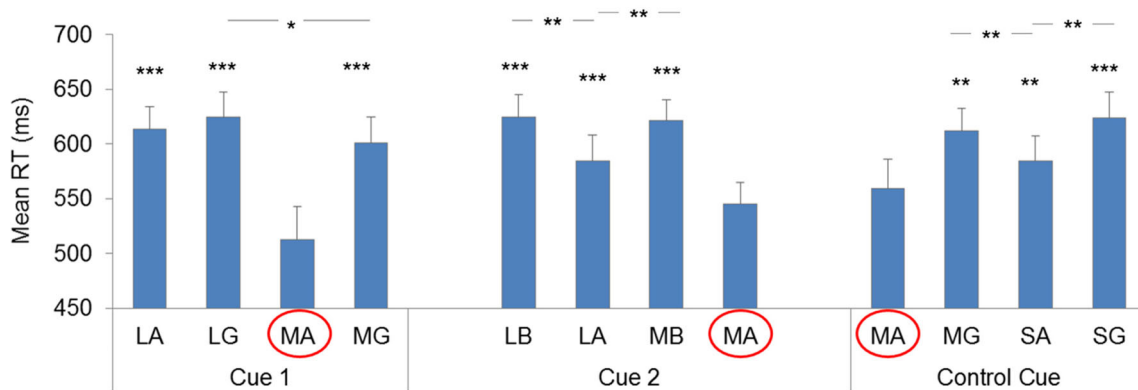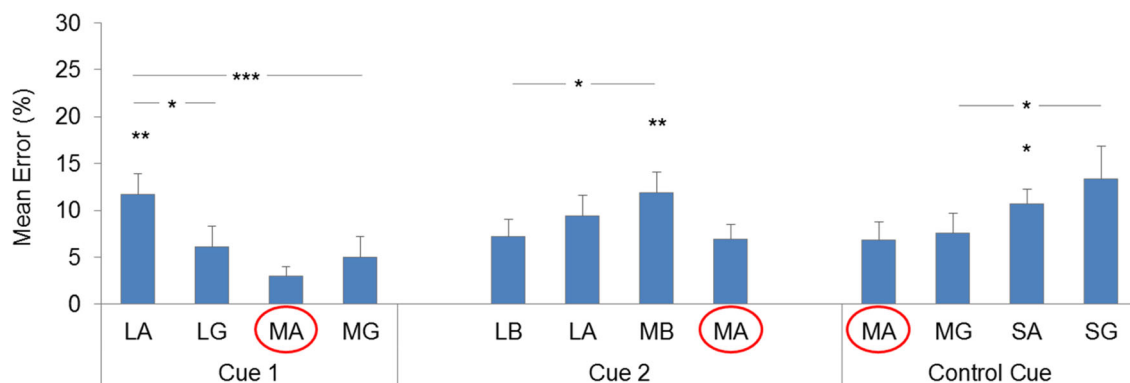
Fig. 5 Predictions for Experiment 4 (top panel), and the results; mean RT (middle panel) and error scores (bottom panel). As predicted, the cues matching both the specific size (medium) and color (aqua) of the target now reliably attracted attention, as reflected in shorter RT and lower errors for the target-matching MA cue. Error bars represent 1 *SEM*. *$p \leq .05$, **$p \leq .01$. ***$p \leq .001$. Asterisks on top of the histograms are for the comparison with the the MA cue that was predicted to capture. (Color figure online)

Apart from the predicted differences, Experiment 4 showed significant differences among the cues that were not predicted to attract attention (e.g., the MG and LG cues in Cue Display 1; see Fig. 5). Inspection of the mean RT indicates that the cues with the target color (LA, SA cues) may have facilitated responding in addition to the target-matching MA cue. However, the mean errors did not show the same results

pattern, and even showed the opposite results in Cue Display 1, where the LA cue produced the highest error rates.

Moreover, across Experiments 1 to 3, the magnitude of cueing effects for each cue display varied rather unsystematically and without indicating any regularity underlying the data. Thus, these results need to be interpreted with caution. Overall, the results of Experiments 1–4, while consistently

showing the predicted results pattern, also tended to show other differences between the cues that were not predicted, which indicates that the present paradigm may be more sensitive to small differences and interactions than the standard paradigm, in which the nontargets and context cues are all identical (e.g., Becker et al., 2013, 2017; Folk & Remington, 1998; Harris et al., 2013; Schönhammer et al., 2016).

General discussion

The present study yielded several interesting results. First, we confirmed that relational search is a kind of "default" search mode, also in conjunction search, as reflected by the observation that observers prefer relational search when both relational and feature-specific search strategies are viable (i.e., when both the relative and exact target features remain constant across trials; see Experiment 1). To date, this has only been shown in one previous study that used data-limited search displays in conjunction search (of 100 ms; Becker et al., 2017). The present study confirms that relational search is the preferred search strategy, also in the more standardly used unlimited search displays (that are visible until the response).

A second finding was that attention could be top down tuned to the specific feature values of a conjunction target to elicit attentional capture for target-matching conjunction cues. As discussed, this is not self-evident, as relational search is the preferred search mode also in conjunction search (Becker et al., 2017), and conditions that require feature-specific search have formerly been shown to lead to performance decrements (Arguin & Saumier, 2000; Bauer et al., 1996; Becker, Harris, Venini, & Retell, 2014b; Brand et al., 2014; D'Zmura, 1991; Hodsoll & Humphreys, 2001). Despite this, the present study shows that feature-specific tuning to the features of a conjunction target promotes efficient search, to the extent that an irrelevant conjunctions cue reliably attracts attention, even if it is completely irrelevant and is presented only briefly (100 ms).

The cue was a genuine feature conjunction stimulus, and it did not have a higher feature contrast than any of the other cues in the display (i.e., it was not salient). Although capture by genuine conjunction cues has been shown in one previous study (Becker et al., 2017), the conjunction displays in the present study were better controlled in terms of saliency, as each cue element was a unique item, and we repeatedly presented the same cue displays across experiments (i.e., Experiments 1 and 2; Experiments 3 and 4), showing that each time, a different cue captured (compare Fig. 2 with Fig. 5). This provides the most direct and compelling evidence for the claim that stimuli can reflexively capture attention by virtue of the top-down search settings, regardless of possible existing differences in bottom-up feature contrasts. This argues against accounts that attribute capture to a large extent to bottom-up

saliency (e.g., Itti & Koch, 2001; Martinez-Trujillo & Treue, 2004; Navalpakkam & Itti, 2006, 2007; Theeuwes, 1992), and/or theories that require the target to have a unique feature to allow efficient search (e.g., Treisman & Gelade, 1980; but see Treisman & Sato, 1990; Wolfe, 1994, 1998).

Arguably the most important finding of the present study was that relational versus feature-specific search modes are determined separately and individually for different stimulus dimensions in conjunction search. In particular, we established that attention can be tuned to the relative target feature in one dimension (e.g., color) and the exact feature value in another dimension (e.g., size; see Experiment 2; or vice versa; see Experiment 3). This renders search strongly context dependent and independent of the exact feature values in one dimension (e.g., selection of bluest), and context-independent but contingent on the exact feature match in another dimension (e.g., selection of a specific, medium size).

One noteworthy implication of these findings is that conjunction search is not based on a holistic target template, which limits selection to objects that match an integrated representation of the target object (e.g., Eimer & Grubert, 2014; see also Kahneman et al., 1983; Treisman et al., 1983). Of note, relational search causes capture by conjunction cues that look completely different from the target. For instance, in Experiment 1, a large, blue cue attracted attention even though observers were looking for an aqua, medium target. The finding that a very target-dissimilar item attracted attention suggests that the cue was not compared with an integrated representation of the target object—at least not prior to selecting it, which means that holistic target representations played little to no role in selection (cf. Becker et al., 2017). Performance in the present conjunction search task was probably also not mediated by other higher-order, conceptual, or semantic processes (Duncan, 1984; Wyble et al., 2013). Instead, the findings provide converging evidence for the view that conjunction search is driven by independent, feature-based attentional mechanisms (e.g., Treisman & Sato, 1990; Wolfe, 1994, 1998).

Previous studies have already shown that attention is typically more strongly biased towards the more discriminable feature of a conjunction target (e.g., more strongly to color than shape, if the target differs more in color; e.g., Hannus, van den Berg, Bekkering, Roerdink, & Cornelissen, 2006; Williams & Reingold, 2001; but see Treisman & Sato, 1990), or towards the feature subgroup that is more likely to contain the target (subset search; e.g., Egeth, Virzi, & Garbart, 1984; Kaptein, Theeuwes, & van der Heijden, 1995), indicating that attention can be tuned independently to different stimulus dimensions. Similarly, recent EEG studies have shown that the amount of attention allocated to a conjunction target can be predicted from a disjunctive search for either feature in isolation (Andersen, Müller & Hillyard, 2015; see also Eimer & Grubert, 2014; Kiss, Grubert, & Eimer, 2013), and that

attention can be allocated to multiple different locations in the display, to speed processing of the corresponding items (Eimer & Grubert, 2014). The present study adds to this body of research by confirming that conjunction search is feature-based rather than object-based, and showing that attention is simultaneously tuned to the different stimulus dimensions of color and size. Of note, in classical conjunction studies (e.g., Becker et al., 2017), the nontargets (and context cues, if applicable) always share one feature with the target. Hence, theoretically, capture by a fully matching cue could be due to tuning attention only to the color or size dimension on a trial-by-trial basis. If, on any given trial, only the bluest *or* the largest items were attended in the cue display, averaging across the trials would show an advantage for the fully matching, largest and bluest item, because it is *always* among the selected items, whereas the partially matching items that match only the relative color or the relative size would only be selected 50% of the time. While the standard conjunction search tasks thus seem consistent with sequential tuning and switching between color and size on a trial-by-trial basis, the present study is inconsistent with this scenario: Of note, Cue Display 3 always contained a small, green (SG) cue that did not match any of the relative or specific features of the target. This SG cue would never be among the selected items, and, thus, in a sequential account, it should result in elevated RTs that differ as much from the partially matching cues (MG, SA cues) as these differ from the fully matching (MA) cue. A corresponding result pattern was observed only in Experiment 2 (see Fig. 3), but not in any of the other experiments (see Figs. 2, 4, 5). This suggests that attention is not sequentially tuned to either color or size on a trial-by-trial basis, but simultaneously to both color and size within the same trial, which provides a clear advantage to the fully matching cues and prevents selection of the partially matching cues.

Moreover, the findings of Experiments 1 to 4 have two other important implications—first, with regard to how feature search and relational search are related to each other, and second, with respect to how control mechanisms for attention are implemented.

With respect to the first question, we can ask how the visual system extracts information about feature relationships. According to traditional accounts of visual search, search is based on feature maps or feature detectors (e.g., Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe, 1994), which are topographically organized sensory neurons that signal the location of a particular feature in the visual field (e.g., red, green, blue), and can be top-down modulated—for instance, by increasing the gain of feature detectors that would respond to the target feature(s) (Wolfe, 1994; but see Treisman & Sato, 1990). The feature detectors are usually modeled such that they respond to a limited range of features and operate independently of each other. The finding that search is often

relational raises the question whether the feature detectors are instead relational, in that they only (or predominantly) signal the presence of relative features (e.g., the presence of a redder object in the center of the receptive field than in the surround; e.g., Becker, 2013b). In this case, information about specific feature values and the corresponding context-independent representation would be derived at a later stage of visual processing. Alternatively, feature detectors could be feature specific, with information about feature relationships being extracted at a later stage.

The present study does not answer the question whether information about feature relationships is extracted at an earlier or later level than feature-specific information, but it suggests a third possibility—namely, that relational and feature-specific search are implemented at the same level. The fact that we never needed to enforce tuning to relative features (e.g., by varying the feature value of the target) suggests that observers may always start out with a relational search strategy, and switch only to feature-specific search when relational search leads to frequent erroneous selection of the nontarget items (e.g., Becker et al., 2017; Becker, Harris, Venini, & Retell, 2014b). Thus, it is possible that requirement to narrowly tune attention to a specific feature value involves inhibition (e.g., of feature maps responding to blue or the bluest item, in search for aqua). Inhibition of feature maps would decrease the target signal and/or increase the noise of the target representation, which could explain why search is typically less efficient in feature-specific search mode (e.g., Becker, 2014; see also Navalpakkam & Itti, 2006, 2007; Wolfe, 1998). However, at this stage, this is speculative and would require further research.

A second interesting implication concerns the implementation of attentional control mechanisms. The present study shows that selection of the search mode for a conjunction target is made on the basis of how the target can be discriminated from the nontarget features within one feature dimension, and largely independently across different feature dimensions. This suggests the existence of local, specialized attentional control mechanisms that operate independently of each other within each stimulus dimension, and adjust the tuning in response to selection feedback.

Previous studies already found evidence for fairly low-level attentional control mechanisms that adjust control settings differently within a given stimulus dimension, according to the affordances of the task. For instance, Harris, Becker, and Remington (2015) showed that attentional capture by color singletons can be due to a "color singleton search mode," which leads to capture by all color singletons but excludes salient stimuli from other stimulus dimensions (e.g., motion singletons; Harris et al., 2015). In addition, there is evidence of dimension-specific control processes that may not only affect early attentional processes, but also later processes concerned with response selection (e.g., Becker, Grubert, & Dux,

2014a; Müller, Heller & Ziegler, 1995). Previous studies on the priming effect already found evidence for fairly low-level attentional control mechanisms that bias attention to different (relative) features in response to the target features selected on the previous trial (Becker, 2010, 2013a; Becker, Valuch, & Ansorge, 2014c; see also Maljkovic & Nakayama, 1994, 1996). In the present study, priming effects were largely absent, which indicates that the control settings mostly did not need to be readjusted on a trial-by-trial basis to guarantee target selection. Yet, collectively, these results suggest the existence of attentional control mechanisms that operate largely independently of each other in different dimensional modules or units, and adjust the gain of different features (or relative features) largely autonomously to optimize performance in the task.

In the past, attention research has often been dominated by the question of whether a given process is top-down or bottom-up (e.g., Folk et al., 1992; Theeuwes, 1992; Yantis, 1993). Adaptations in search settings that occur in these local units can neither be classified as top down nor bottom up, as they most likely reflect a relatively restricted, local feedback loop. Logic dictates that the first half of a feedback loop contains bottom-up processing (with information travelling from lower-level visual areas to higher-level areas), whereas the second half of a feedback loop contain top-down processing (with the reverse information flow). Hence, in addition to top-down and bottom-up processing, we need to acknowledge adjustments of search settings made in response to local feedback, which cannot be described in terms of top-down or bottom-up processes. Current models of attention need to acknowledge these local feedback control processes in addition to the known top-down and bottom-up control mechanisms, and integrate them into the existing architecture. The current evidence suggests that local feedback mechanisms operate largely autonomously and automatically across different stimulus dimensions, but are contingent on higher-order, top-down goals (e.g., Becker, 2007, 2008a, 2008b; Fecteau, 2007).

The terms of "top-down" versus "bottom-up" processing are not strictly limited to the flow of information (from higher-order to lower-level areas or vice versa), but are sometimes used as shorthand to refer to internal, user-driven, and voluntary processing (top-down) versus external, stimulus-driven, and involuntary processing (bottom-up; e.g., Wolfe, 1994, 1998). Here, it is important to acknowledge that attention is, to a large extent, controlled by internal, user-driven control settings that condition otherwise automatic or "involuntary" processes ("contingent automaticity"; e.g., Bargh, 1989, 1992; Becker, 2007; 'contingent capture'; Folk & Remington, 1998; see also Anderson & Folk, 2014). The present study found evidence that such control settings are instantiated through task-driven adaptation fed by local feedback systems that determine search modes within a given stimulus dimension without affecting the search setting in another stimulus dimension. It is likely that these adjustments were largely automatic, and proceeded without conscious knowledge or assistance from voluntary acts, in response to local feedback and perhaps, as part of a trial-and-error strategy that only requires observers to engage in the task. With this, the mechanism in question probably exemplifies an instance of *contingent automaticity*, where the continued execution of a largely reactive, feedback-driven system depends mainly on observers continuing to do the task, without further assistance from conscious, voluntary processes. It is possible that the majority of visual search mechanisms rely on semiautomatic processes and local control mechanisms that are feedback driven. Further research is required to describe these semiautomatic processes in more detail and clarify to what extent they are top-down contingent and top-down penetrable.

References

Andersen, S. K., Müller, M. M., & , S. A. (2015). Attentional selection of feature conjunctions is accomplished by parallel and independent selection of single features. *Journal of Neuroscience*, *35*(27), 9912–9919. doi:https://doi.org/10.1523/JNEUROSCI.5268-14.2015

Anderson, B. V. A., & Folk, C. V. L. (2014). Conditional automaticity in response selection: Contingent involuntary response inhibition with varied stimulus-to-response mappings. *Psychological Science*, *25*, 547–554.

Arguin, M., & Saumier, D. (2000). Conjunction and linear nonseparability effects in visual shape encoding. *Vision Research*, *40*, 3099–3115.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. Perception and Psychophysics, 55, 485–496.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. A. Bargh & J. S. Ulman (Eds.), Unintended thought (pp. 3–51). New York, NY: Guilford Press.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology*, *105*, 181–199.

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Visual search for color targets that are or are not linearly separable from distractors. *Vision Research*, *36*, 1439–1465.

Becker, S. I. (2007). Irrelevant singletons in pop-out search: Attentional capture or filtering costs? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 764-787.

Becker, S. I. (2008a). Can intertial effects of features and dimensions be explained by a single theory? *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1417–1440.

Becker, S. I. (2008b). The mechanism of priming: Episodic retrieval or priming of pop-out? *Acta Psychologica*, *127*, 324–339.

Becker, S. I. (2010). The role of target-distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology. General*, *139*(2), 247–265. doi:https://doi.org/10.1037/a0018808

Becker, S. I. (2013a). Simply shapely: Relative, not absolute shapes are primed in pop-out search. *Attention, Perception, & Psychophysics*, *75*, 845–861.

Becker, S. I. (2013b). Why you cannot map attention: A relational theory of attention and eye movements. *Australian Psychologist*, *48*, 389–398.

Becker, S. I., Folk, C. L., & Remington, R. W. (2010). The role of relational information in contingent capture. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1460–1476. doi:https://doi.org/10.1037/a0020370

Becker, S. I., Folk, C. L., & Remington, R. W. (2013). Attentional capture does not depend on feature similarity, but on target-nontarget relations. *Psychological Science*, *24*(5), 634–47. doi:https://doi.org/10.1177/0956797612458528

Becker, S. I., Grubert, A., & Dux, P. E. (2014a). Distinct neural networks for target feature versus dimension changes in visual search, as revealed by EEG and fMRI. *NeuroImage*, *102*, 798–808.

Becker, S. I., Harris, A. M., Venini, D., & Retell, J. D. (2014b). Visual search for color and shape: When is gaze guided by feature relationships, when by feature values? *Journal of Experimental Psychology. Human Perception and Performance*, *40*(1), 264–91. doi:https://doi.org/10.1037/a0033489

Becker, S. I., Harris, A. M., York, A., & Choi, J. (2017). Conjunction search is relational: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1828–1842.

Becker, S. I., Valuch, C., & Ansorge, U. (2014c). Color priming in pop-out search depends on the relative color of the target. *Frontiers in Psychology, 5*(289), 1–11.

Brand, J., Oriet, C., Johnson, A.P., & Wolfe, J.M. (2014). Flexible cue combination in the guidance of visual search. *Acta Psychologica*, *153*, 129–138.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*, 523-547.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.

D'Zmura, M. (1991). Color in visual search. *Vision Research*, *31*, 951–966. doi:https://doi.org/10.1016/0042-6989(91)90203-H

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. doi:https://doi.org/10.1037/0096-1523.10.1.32

Eimer, M. (1996). The N2pc as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*, 225–234.

Eimer, M., & Grubert, A. (2014). The gradual emergence of spatially selective target processing in visual search: From feature-specific to object-based attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1819–1831.

Fecteau, J. H. (2007). Priming of pop-out depends on the current goals of observers. *Journal of Vision*, *7*(6), 1–11.

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 847–858. doi:https://doi.org/10.1037/0096-1523.24.3.847

Folk, C. L., & Remington, R. W. (2008). Bottom-up priming of top-down attentional control settings. *Visual Cognition*, *16*, 215–231. doi:https://doi.org/10.1080/13506280701458804.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030–1044. doi:https://doi.org/10.1037/0096-1523.18.4.1030

Geng, J. J., DiQuattro, N. E., & Helm, J. (2017). Distractor probability changes the shape of the attentional template. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1993–2007.

Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set for displaywide visual features. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 699–706.

Hannus, A. A., van den Berg, R., Bekkering, H., Roerdink, J. B., & Cornelissen, F. W. (2006). Visual search near threshold: Some features are more equal than others. *Journal of Vision*, *6*(15), 523–540.

Harris, A. M., Becker, S. I., & Remington, R. W. (2015). Capture by color: Evidence for dimension-specific singleton capture. *Attention, Perception, & Psychophysics*, *77*, 2305–2321.

Harris, A., Remington, R., & Becker, S. I. (2013). Feature specificity in attentional capture by size and color. *Journal of Vision*, *13*(2013), 1–15. doi:https://doi.org/10.1167/13.3.12.doi

Hodsoll, J., & Humphreys, G. (2001). Driving attention with the top-down: The relative contribution of target templates to the linear separability effect in the size dimension. *Perception & Psychophysics*, *63*, 918–926.

Irons, J. L., & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics, 78*(7), 2031–2048.

Itti, L., & Koch, C. (2001) Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203.

Kahneman, D., Treisman, A., & Burkell, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 510–522.

Kaptein, N. A., Theeuwes, J., & van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1053–1069. doi:https://doi.org/10.1037/0096-1523.21.5.1053

Kiss, M., Grubert, A., & Eimer, M. (2013). Top-down task sets for combined features: Behavioral and electrophysiological evidence for two stages in attentional object selection. *Attention, Perception, & Psychophysics*, *75*, 216 –228. doi:https://doi.org/10.3758/s13414-012-0391-z

Koivisto, M., & Revonsuo, A. (2007). How meaning shapes seeing. *Psychological Science*, *18*, 845–849.

Leonard, C. J., & Egeth, H. E. (2008). Attentional guidance in singleton search: An examination of top-down, bottom-up, and intertrial factors. *Visual Cognition*, *16*, 1078–1091.

Ludwig, C. J. H., & Gilchrist, I. D. (2002). Stimulus-driven and goal-driven control over visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 902–912. doi:https://doi.org/10.1037/0096-1523.28.4.902

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22,* 657– 672. doi:https://doi.org/10.3758/BF03209251

Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: The role of position. *Perception and Psychophysics, 58,* 977–991.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, *14*(1), 744–751. doi:https://doi.org/10.1016/j.cub.2004.04.028

Meeter, M., & Olivers, C. N. L. (2014). Target features and target-distractor relation are both primed in visual search. *Attention, Perception, & Psychophysics*, *76*, 682–694.

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, *57*, 1–17.

Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron*, *53*(4), 605–617. doi:https://doi.org/10.1016/j.neuron.2007.01.018

Navalpakkam, V., & Itti, L. (2006). Top-down attention selection is fine grained. *Journal of Vision*, *6*, 1180–1193. doi:https://doi.org/10.1167/6.11.4

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Scolari, M., & Serences, J. T. (2009). Adaptive allocation of attentional gain. *Journal of Neuroscience, 29,* 11933–11942.

Scolari, M., Byers, A., & Serences, J. T. (2012). Optimal deployment of attentional gain during fine discriminations. Journal of Neuroscience, 32, 7723-7733.

Schönhammer, J. G., Grubert, A., Kerzel, D., & Becker, S.I. (2016). Attentional guidance by relative features: Behavioral and electrophysiological evidence. *Psychophysiology*, *53*, 1074–1083.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 459–478. doi:https://doi.org/10.1037/0096-1523.16.3.459

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics*, *33*, 526–532.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, *123*, 146–160.

Williams, D. E., & Reingold, E. M. (2001). Preattentive guidance of eye movements during triple conjunction search tasks: The effects of feature discriminability and saccadic amplitude. *Psychonomic Bulletin & Review*, *8*, 476–488.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. doi:https://doi.org/10.3758/BF03200774

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), Attention (pp. 30–73). London, UK: University College London Press.

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Journal of Vision*, *37*, 25–43.

Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 861–871. doi:https://doi.org/10.1037/a0030517

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(3), 676–681. doi:https://doi.org/10.1037/0096-1523.19.3.676

Yu, X., & Geng, J. J. (2019). The attentional template is shifted and asymmetrically sharpened by distractor context. *Journal of Experimental Psychology: Human Perception and Performance*, *45,* 336–353.