Conjunction Search Is Relational: Behavioral and Electrophysiological Evidence

Stefanie I. Becker, Anthony M. Harris, Ashley York, and Jessica Choi
The University of Queensland

Attention selects behaviorally relevant stimuli for further capacity-limited processing and gates their access to awareness. Given the importance of attention for conscious perception, it is important to determine the factors and mechanisms that drive attention. A widespread view is that attention is biased to the specific feature values of a conjunction target (e.g., vertical, red, medium). By contrast, the results of the present study show that attention is tuned to the 2 relative features that distinguish a conjunction target from the irrelevant nontargets (e.g., larger and bluer). Moreover, an irrelevant conjunction cue that is briefly presented prior to the target can automatically attract attention, even in the absence of any feature contrasts. Importantly, automatic orienting to the conjunction cue was completely independent of the physical similarity between cue and target, and depended only on whether the conjunction cue matched the relative features of the target. These results demonstrate that attentional orienting is determined by a mechanism that can rapidly extract information about feature relationships and guide attention to the stimulus that best matches the relative attributes of the target. These results are difficult to reconcile with extant feature-specific accounts or object-based accounts of attention and argue for a relational account of conjunction search.

Public Significance Statement

How do we select important information from cluttered visual scenes? As previous research mainly used simple stimuli to examine how we allocate attention, it is unknown how we attend to more complex stimuli, such as feature conjunctions. Among the discussed possibilities are that attention is biased toward (a) an entire object, (b) its constituent features such as its color and size, or (c) its context-dependent attributes, which entails processing of all features and biasing attention to the relative features of an object (e.g., larger and redder). To distinguish between these theories, we measured participants' behavior and an electroencephalographic (EEG) marker of attention (N2pc) in response to irrelevant conjunction stimuli that matched 1 or more of the target features versus only its relative features. The results showed that attention was biased in a context-dependent manner to the target, demonstrating that context-dependent search strategies are also applied in search among more complex stimuli.

Keywords: visual search, attentional capture, conjunction cue, context, relational account

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 and gates access to awareness. Hence, it is important to understand what factors control and guide visual attention. Previous research has shown that attention can be guided both by the goals of the observer in a top-down manner (e.g., Folk & Remington, 1998) as

well as by salient stimuli in a bottom-up, stimulus-driven manner (e.g., Itti & Koch, 2000; Li, 2002). For instance, in search for a specific target such as a red item, attention will be more strongly attracted to other red items, including when they are irrelevant, than differently colored items (e.g., green items; Folk & Remington, 1998; Ludwig & Gilchrist, 2002). This *similarity effect* has been replicated in numerous studies and shows that we can actively tune attention to elementary features (e.g., colors), which guides attention to locations that contain corresponding objects.

Different tuning mechanisms have been proposed to account for top-down effects on selection (e.g., Duncan & Humphreys, 1989; Folk, Remington, & Johnston, 1992; Folk & Remington, 1998; Maunsell & Treue, 2006; Navalpakkam & Itti, 2006, 2007; Treisman & Sato, 1990; Wolfe, 1994). One of the most wide-spread views is the *feature-similarity view*, which proposes that attention is biased to the feature value of a sought-after target (e.g., Martinez-Trujillo & Treue, 2004). Other accounts propose that we can alternatively or additionally inhibit features that are known to

be task-irrelevant (e.g., Treisman & Sato, 1990). According to *optimal tuning accounts*, attention is usually tuned to the feature that optimally distinguishes the target from the nontargets, which can deviate from the target feature (Lee, Itti, Koch, & Braun, 1999; Navalpakkam & Itti, 2007; Scolari & Serences, 2010). Especially when the target is similar to the nontargets, tuning attention to an "exaggerated" target feature that is shifted away from the nontarget feature (e.g., tuning to 60° in search for a 55° target) can be more optimal, as it increases the signal-to-noise ratio for the target feature. Although current theories vary in their predictions about which feature we will tune attention to, they concur in that attention should be most strongly attracted to items that are maximally similar to the tuned-to feature (cf. Becker, 2010).

Contrary to this view, Becker (2010) proposed that attention can be biased to feature relationships or relative features such as the relative color or size of the target (e.g., redder, larger, darker). According to this *relational account*, biasing attention to a relative feature of the target guides attention to items that have the same relative features, regardless of the physical similarity to the target. In line with the relational account, Becker, Folk, and Remington (2013) found that target-similar items can fail to attract attention if they fail to match the target's relative feature, whereas target-dissimilar items can strongly attract attention, provided that they have the same relative features as the target. For instance, in one study, participants had to search for an orange target among three yellow-orange nontargets (i.e., redder target) while ignoring an irrelevant salient cue that was briefly flashed prior to the target display (see also Becker, Folk, & Remington, 2010). The cue was presented among three other cues (which constituted the cue context), and the colors varied such 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-similar orange cue failed to attract attention when it was embedded among other red cues in the cueing display (yellower cue). In turn, a target-dissimilar cue that matched the targets' relative feature (redder) reliably attracted attention, including when it had the same color as the nontargets (yellow-orange cue in yellow cue context) or when it had to compete against other cues that had the target color (e.g., red cue in orange cue context; Becker et al., 2013). These results were inconsistent with current feature-based accounts of attention, and provided strong evidence for a relational account.

Subsequent studies, moreover, showed that relational search is applied in a variety of search tasks, including in search for size, shape and luminance targets, and in pop-out displays (i.e., when the nontargets are all identical) as well as feature search displays (i.e., when the nontargets all have different features; Becker, 2013; Becker, Harris, Venini, & Retell, 2014). Top-down tuning to the exact feature value of the target can also be observed, but has to be enforced by rendering a relational search strategy impossible—for instance, by randomly varying the nontarget features such that the target is not reliably the largest/smallest, darkest/lightest, or reddest/yellowest item (e.g., presenting an orange target randomly among all-red or all-yellow nontargets; Harris, Remington, & Becker, 2013). In these instances, attention is indeed biased to the exact feature value of the target (e.g., medium or orange). Critically, though, when the target and nontarget features remain con-

stant, attention is biased to the relative feature of the target, indicating that relational search is a *default* search strategy.

However, relational search is also subject to several limitations: First, tuning attention to the relative feature of a sought-after item requires prior knowledge or expectations about how the target will differ from the irrelevant items in the context. Such knowledge or expectations could result from extensive experience, or a fast "feature averaging process" that allows immediate extraction of the dominant feature in unfamiliar scenes (e.g., Chong & Treisman, 2005). Second, relational search can only optimally support target selection when the target has a relatively extreme feature (e.g., largest, reddest item), such that it differs in a single direction from the nontarget items, as tuning to redder will result in selection of the reddest item in the visual field (e.g., Becker, Grubert, & Dux, 2014; but see Becker et al., 2013).

In line with this limitation, it is known that search is much more efficient when the target has a relatively extreme feature (e.g., largest, reddest, darkest) than when it has an intermediate feature (e.g., medium target among small and large nontargets). This *linear separability effect* has previously been attributed to bottom-up processes (e.g., D'Zmura, 1991) or dissimilarity among the nontargets (e.g., large, small; Bauer, Jolicoeur, & Cowan, 1996; Hodsoll & Humphreys, 2001). However, in line with the relational account, Brand, Oriet, Johnson, and Wolfe (2014) recently found that the linear separability effect is indeed due to top-down tuning to relative features.

Critically, Brand, et al. (2014) pointed to a third possible limitation of relational search, in that they claimed that *conjunction search* may be feature-specific, not relational. In a conjunction search task, the target differs only in a combination of two or more features from the nontargets, such as the particular color and orientation of the target (e.g., red vertical target among horizontal red and vertical green nontargets; Treisman & Gelade, 1980; Wolfe, 1994). If conjunction search is indeed always feature-specific, this would impose severe limitations on the relational account, as most searches in the natural environment involve feature conjunctions (e.g., finding a car in a crowded parking lot probably requires tuning attention to the color and shape of the car; Kiss, Grubert, & Eimer, 2013; Wolfe, 1998). So far, it has never been tested whether conjunction search is indeed feature-specific or relational. However, if relational search can only be applied when the target has a relatively extreme feature (e.g., reddest, largest), it logically follows that conjunction search cannot be relational, as a conjunction target always has the same features as the nontargets, and thus possesses no feature more extreme than those of the distractors. This limitation would also be in line with previous formulations of the relational account (e.g., Becker, 2010), and is in line with the currently available evidence on the relational account.

However, an alternative interpretation of the relational account and the previous evidence is that relational search does not require the target to possess a unique extreme feature. Rather, relational search may only require the target to be uniquely defined by a combination of one or more extreme features, such that it is *one of* the largest (or smallest) items in the display. In other words, relational search may be successfully applied as long as the target is not surrounded by more extreme nontargets (that cause the target to differ in directly opposite relative features). In this case, conjunction searches could still be relational, as the target typically

differs consistently in a relative feature from each of the nontargets (e.g., the target is among the reddest and steepest items across all trials).

Moreover, on the basis of the most prominent feature-based accounts of visual search, it seems difficult to reject the possibility of relational conjunction search. For instance, according to the *guided search model* (Wolfe, 1994), conjunction search differs only from feature search in that it requires combining information from multiple stimulus dimensions (e.g., color and size). This integration, however, occurs automatically, as the output from feature-based detectors is summed at the level of the saliency map (e.g., Wolfe, 1994). Given that this integration requires no additional processes, nothing would seem to dictate that attention ought to be tuned to specific feature values (e.g., red and vertical) rather than relative features (e.g., redder and steeper).

Tuning to relative features may, however, seem unlikely according to two other classes of theories. *Late selection views* typically assume that search items can be fully identified before we focus attention on them, and *object-based accounts* of attention typically assume that attention can be tuned to an entire object or all features of an object in a holistic fashion (Kahneman & Treismann, 1984; see also Duncan, 1984; Vecera & Farah, 1994; Wyble & Swan, 2015; Wolfe & Bennett, 1997; Wyble, Folk, & Potter, 2013). If items can be identified prior to selecting them or if attention can indeed be effectively biased to an integrated target representation, selection should be limited to those objects that perfectly match the target features, and selection would therefore be independent of other features in the context (i.e., not relational).

To date, there is also some evidence that we can strategically tune attention to higher order properties or an integrated target representation (e.g., Vecera & Farah, 1994; Wyble et al., 2013; Wyble & Swan, 2015). Still, in classical conjunction search tasks, attention seems to be mainly driven by independent, feature-based attentional mechanisms rather than object-based mechanisms. For instance, several studies found that attention is often more strongly biased toward the more discriminable feature of a conjunction target (e.g., more strongly to color than shape, if the target differs more in color; Egeth, Virzi, & Garbart, 1984; Kaptein, Theeuwes, & van der Heijden, 1995), indicating that attention can be tuned independently to features from different stimulus dimensions. Similarly, recent electroencephalographic (EEG) studies have shown that an irrelevant distractor that shares only one feature with the target can still elicit the electrophysiological marker for attentional capture (N2pc; Eimer & Grubert, 2014), and 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). These findings are difficult to reconcile with a holistic or object-based view of conjunction search and indicate that features are separable at the level of top-down tuning of attention, in line with guided search (Wolfe, 1994).

It is, however, also noteworthy that Eimer and Grubert (2014) recently proposed a *two-stage model of conjunction search*, in which early, feature-based processes are followed by later, object-based processes that operate on the integrated features of a conjunction target (see also Nako, Grubert, & Eimer, 2016). According to this model, initial feature-based processes lead to simultaneous selection of all objects that share one of the target features. This results in multiple attentional selection foci in the visual field at an early stage of visual search (i.e., ~200 to 250 ms

after display onset). True selection of the conjunction target within a single focus of attention (i.e., under exclusion of the nontargets) is only achieved at a later stage, when object-based processes that operate on an integrated target representation determine selection (i.e., ~250 to 300 ms after display onset). According to this two-stage account of conjunction search, it is possible that early selection depends only on a relative match of target and nontarget features, whereas later processes determine selection based on a match in the exact feature values of target and distractor.

Aim of the Present Study

The aim of the present study was to test whether conjunction search is based a relational or feature-specific search strategy. Across four experiments, participants had to search either for a particular color-size conjunction target or a luminance-size conjunction target, and to ignore an irrelevant conjunction cue that was briefly presented prior to the target display. Across seven different cue displays, the features of the conjunction cue were systematically varied, and we critically assessed whether an irrelevant conjunction cue would attract attention depending on a match in the relative features of cue and target, or their physical features.

The first experiments (Experiments 1 and 2) critically tested a relational account of conjunction search against the prevalent feature-specific accounts, which predict capture by target-similar conjunction cues. Subsequent experiments (Experiments 3 and 4) tested the relational account against the optimal tuning account, which allows tuning to an exaggerated target feature that is shifted away from the nontarget features.

To anticipate the results, attentional capture by the conjunction cue depended strongly on the cue context, and was entirely independent on whether the cue matched or mismatched the physical features of the target. These results argue against current feature-based accounts of attention and support a relational account of conjunction search. In Experiment 4, we assessed the two-stage account of Eimer and Grubert (2014), by measuring the EEG of participants to obtain fine-grained information about the time course of attentional capture by the conjunction cue. The results showed that a conjunction cue that matched the relative features of the target attracted attention at an early stage of visual search (150 to 200 ms after display onset). There was no evidence that attention was oriented to perfectly matching cues at a later stage of visual search, and thus no evidence for an additional, object-based attentional mechanism modulating selection at a later stage of visual processing.

Experiments 1 and 2

The first two experiments critically tested a relational account of conjunction search against the prevalent feature similarity views. In Experiment 1, the target was defined by a color-size conjunction, and in Experiment 2, by a luminance-size conjunction. Stimuli could have one of three different (equiluminant) colors that varied from blue to green (blue, aqua, or green), grayscale values varying from black to light gray and were either small (S), medium (M), or large (L). The color (or grayscale value) and size of the target and the irrelevant nontarget items in the target display always remained constant within a condition (e.g., medium, aqua

target among medium green and small aqua nontargets), so that the target could be equally localized by tuning attention to its specific feature values (medium, aqua) or its relative features (bluest, largest). The target display consisted of bars that were tilted to the right or left, and participants had to respond to the orientation of the conjunction target (right, left) with a button press.

The cues consisted of two tilted bars presented to the left and right of the placeholders (see Figure 1A), one of which was always tilted to the left and the other to the right, to avoid priming a particular response. All cue displays contained five cues—two pairs of cues that were always identical, and one cue that that consisted of a unique combination of these features (i.e., the cue was a genuine feature conjunction stimulus). Across different cue displays, the similarity of this conjunction cue to the target was systematically varied, and attentional capture by the cue was inferred when responses to the target were faster and/or more accurate when the cue was presented at the target location (valid trial) than at a nontarget location (invalid trial).

Systematically varying the similarity of the conjunction cue to the conjunction target resulted in seven different cue displays (see Figure 1A and B for an overview). In the full-match condition, the cue display was identical to the target display, namely, the conjunction cue had the same features as the target, and the cue context comprised the same features as the nontarget context (e.g., medium aqua conjunction cue among medium green cues and small aqua cues). In three different relative-match conditions, the conjunction cue always had the same relative color and size as the target (e.g., the cue was among the largest/bluest items in the cueing display), and it either matched only the physical size of the target (size-match condition) or only its color (color-match condition), or the cue matched neither color nor the size of the target (no-match condition). Conversely, in two feature-match conditions, the cue was identical to the target (e.g., medium aqua), but failed to match either the relative color of the target (e.g., cue not bluest: medium aqua cue among medium blue cues and small aqua cues; RelCol–) or its relative size (e.g., cue not largest: medium
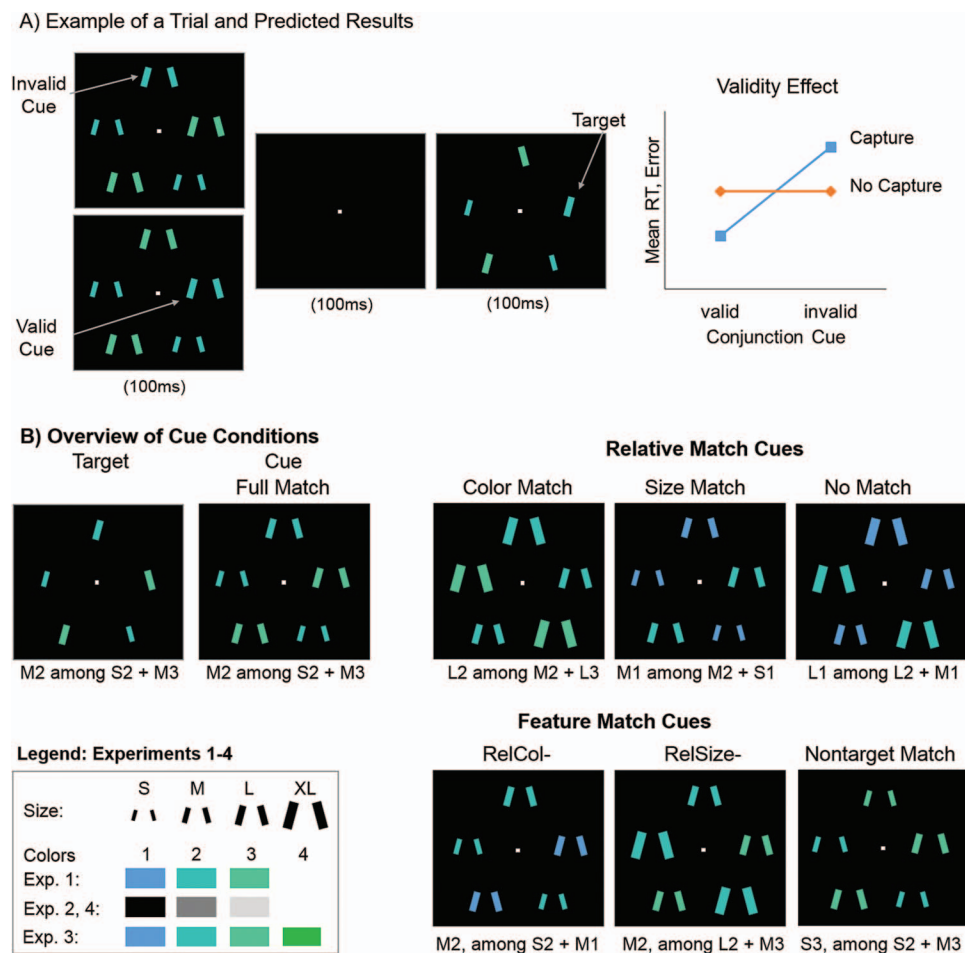

Figure 1. (A) Example of the sequence of events in a single trial. (B) Overview of the seven different cue displays used in Experiment 2. In the figure, the conjunction cue is always displayed on top, whereas its position varied randomly in the experiments. Stimuli are specified by the combination of their color or luminance (Numbers 1–4) and size (S, M, L, XL: small, medium, large, extra-large), with the conjunction cue-target specified first, and the context items (nontarget or cue context) second; for example, M2 among S2 + M3 refers to a medium aqua conjunction cue or target presented among two small aqua and two medium green context items. Deviating from the figure, the background was white.

aqua cue among medium green and large aqua cues; RelSize–). In the final, nontarget-match condition, the four context cues were identical to the four nontarget items in the target display (e.g., medium green and small aqua), but the conjunction cue matched neither the relative nor physical features of the target (small, green cue).

According to the relational account, attention should be tuned to the relative features of the conjunction target (bluer, larger), and, consequently, all cues that are among the largest/bluest items in their respective cue displays should attract attention—regardless of whether they match the target features or not. Hence, the cues in the full-match and the three relative-match conditions should all attract attention (and to the same extent), whereas the target-similar cues in the feature-match conditions and the cue in the nontarget-match condition should all fail to capture. According to a feature-similarity account, capture by the conjunction cue depends on cue-target similarity and is independent of the context. Hence, the conjunction cues in the full-match and the two feature-match conditions should attract attention, as these cues were all identical to the target. In turn, the conjunction cues in the three relative-match and the nontarget-match conditions should all fail to attract attention, as they mismatched one or both of the target features. A corresponding outcome would also be consistent with object-based accounts of attention, which would deny capture by relatively matching cues, as these constitute different objects (e.g., in search for a medium aqua target, a blue large cue could not attract attention).

Another possible, and categorically different, outcome is that conjunction cues will not attract attention at all in the spatial cueing paradigm. As noted above, most theories assume that visual selection results from an interplay of bottom-up, saliency-based, and top-down feature-specific selection mechanisms (e.g., Martinez-Trujillo & Treue, 2004; Navalpakkam & Itti, 2007; Wolfe, 1994). In conjunction search, the target is never salient, and in the present task, the cues were also genuine feature conjunctions that only differed in a conjunction of features from the other cues and were therefore nonsalient. Theories that attribute attentional capture mainly to bottom-up saliency (e.g., Theeuwes, 2010), or that reserve an important role for bottom-up processes in attention (e.g., Li, 2002; Martinez-Trujillo & Treue, 2004), would therefore probably predict that a task-irrelevant conjunction cue would not attract attention.

Method

Participants.   Data were collected from 32 naïve participants, half of which participated in Experiment 1 and half in Experiment 2. Participants with error scores above 40% in any of the conditions were excluded from all data analyses, which led to the exclusion of two participants in Experiment 1, and three participants in Experiment 2.

Apparatus.   A BenQ 19-in. LCD color monitor (resolution: 1,280 × 1,024 pixels; refresh rate: 75Hz) and a Dell Optiplex 745 computer running Presentation software (Neurobehavioral Systems, Berkeley, CA) were used to control the experiment and display the stimuli. A video-based infrared eye tracker (Eyelink 1000, SR Research, Ontario, Canada) recorded eye movements, with a spatial resolution of 0.1° and a temporal resolution of 500 Hz.

Stimuli and design.   Stimuli consisted of small (1.0° × 0.3°), medium (1.7° × 0.4°) or large (2.3° × 0.6°) bars that were tilted 15° to the right or left, and were presented on the inside or outside of gray outlined placeholder boxes (2.4° × 2.4°), 7.8° from the center of the display. In Experiment 1, we used equiluminant stimuli that differed in color and that could be either blue (*xy*: .182, .212), bluish green (*xy*: .203, .288), or green (*xy*: .227, .374). In Experiment 2, grayscale stimuli were used that varied in luminance (dark gray: 23.5 cd/m²; gray: 52.3 cd/m²; light gray: 88.8 cd/m²). To avoid that effects can be attributed to the specific color-size combinations of the cues and targets, the assignment of colors (luminance) varied across participants: In Experiment 1, half of the participants searched for a medium, aqua target among medium green and small aqua nontargets (target larger/bluer), whereas the other half searched for the same target among medium blue and small aqua nontargets (target larger/greener). Correspondingly, half of the participants in Experiment 2 searched for a medium gray bar among medium light gray and small gray bars (target larger/darker), whereas the other half searched for a target embedded among dark-gray medium bars and gray, small bars (target larger/lighter). The colors or gray-scale values of the cues were recoded accordingly, so that, for instance, in the no-match condition, the conjunction cue was either green or blue, depending on the condition (for an overview of the conditions, see Figure 1). The target was always a medium (larger) item across all conditions, because pilot tests had shown that search for a smaller target led to high error scores (but see Experiment 4).

The number of right- and left-tilted items in the target and cue display was controlled such that each display contained an equal number of right- and left-tilted items (exempting the target). Moreover, the positions of the conjunction cue and target were controlled such the target was presented an equal number of times at each location, and the conjunction cue coincided with the target location on 50% of all trials. We opted for this paradigm rather than rendering the cue completely nonpredictive, because previous studies have shown no effect of cue predictiveness on capture (e.g., Folk et al., 1992), and the present paradigm allowed testing all of the critical conditions in a maximally efficient manner (but see Experiment 4 for results with nonpredictive cues). Participants completed 560 trials in each experiment, which yielded 40 trials for each cue and validity condition.

Procedure.   Participants were seated in a normally lit room, with their chin and forehead resting on the chin- and forehead rests of the eye tracker, and viewed the screen with a monitor-head distance of 62 cm. Prior to the experiment, participants were shown an example of the target display and instructed to report whether the target was tilted to the right or left by pressing the right or left mouse button. Moreover, to avoid that the results were contaminated by eye movements, participants were instructed to maintain fixation on the central fixation cross throughout the entire trial.

Each trial started with the presentation of the fixation cross and the cueing display was presented after a random duration of 500 to 1,000 ms (contingent upon participants fixating on the cross). The cueing display (100 ms) was followed by a fixation display (100 ms) and the target display (100 ms), followed by a blank display that was presented until response. After the response, participants were provided with a feedback display containing the words "Correct" (500 ms) or "Wrong" (1,200 ms), plus, optionally, the words

"Fixation not detected" (1,200 ms), when the gaze was not continuously within 50 pixels of the central fixation cross during the trial. The next trial started after presentation of a blank display (250 ms), with the presentation of the fixation cross.

Results

Data. Trials with eye movements and trials with either short (<100 ms) or long (>1,500 ms) reaction times (RTs) were excluded from all analyses, which led to a loss of 0.85% of all data in Experiment 1, and 2.76% of all data in Experiment 2. The main variables of interest were the mean RTs for each condition, which are depicted in Figure 2A. The mean error scores did not show any indication of speed–accuracy trade-offs in any of the conditions or experiments (see Figure 2B).

Experiment 1. Mean correct RTs were statistically analyzed with a 7 × 2 within-subjects ANOVA comprising the variables conjunction cue condition (1–7) and cue validity (valid, in-

valid). For Experiment 1, the results showed a significant main effects of the cue condition, $F(6, 78) = 11.93$, $p < .001$, $\eta^2 = .48$, cue validity, $F(1, 13) = 34.59$, $p < .001$, $\eta^2 = .73$, and a significant Cue × Validity interaction, $F(6, 78) = 17.89$, $p < .001$, $\eta^2 = .58$. The interaction was due to the fact that validity effects occurred only in a subset of cue conditions (see Figure 2A, red line graph).

Paired, two-tailed t tests showed that the conjunction cues in the full-match condition and the relative-match conditions (size match, color match, no match) all attracted attention, as reflected by significant validity effects (full match: 65 ms; $t[13] = 4.42$, $p = .001$; color match: 68 ms; $t[13] = 4.16$, $p = .001$; size match: 58 ms; $t[13] = 4.00$, $p = .002$; no match: 109 ms; $t[13] = 6.54$, $p < .001$). By contrast, the conjunction cue failed to attract attention in the two feature-match conditions, in which the cue was identical to the target. Similarly, the cue failed to capture in the nontarget-match condition, in which the
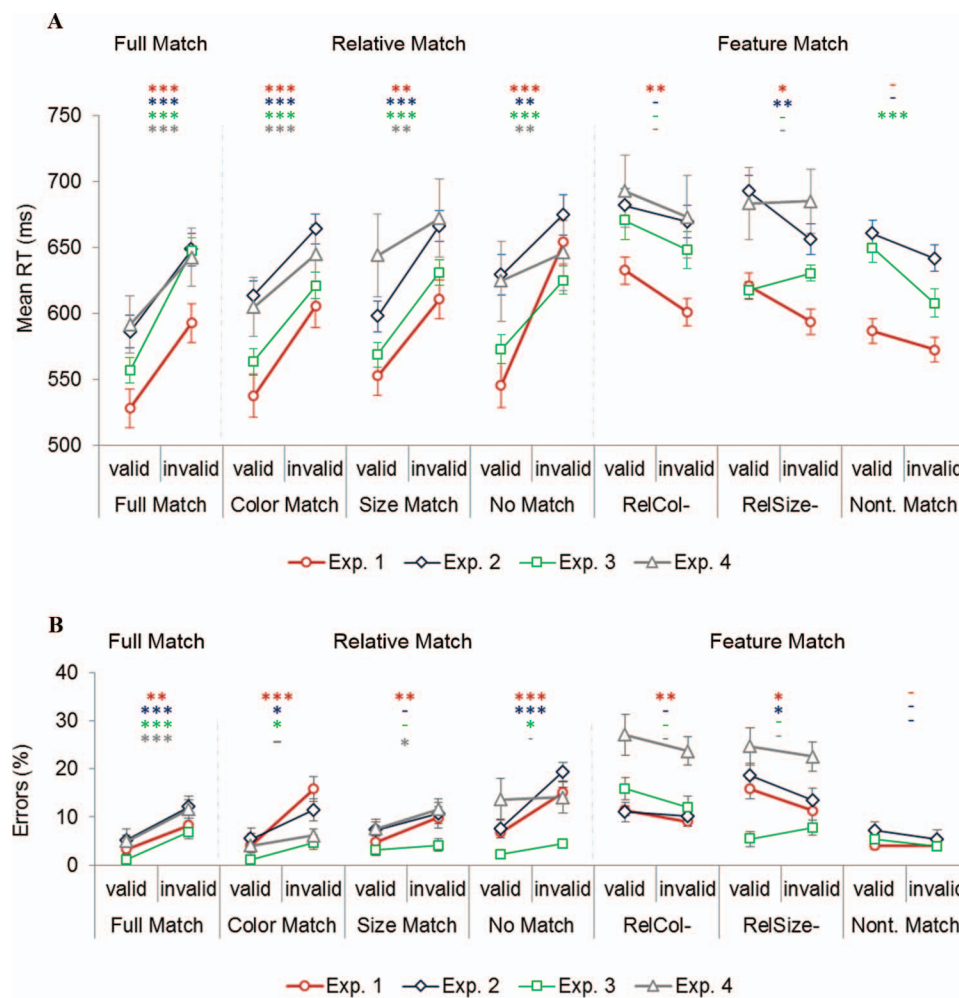

Figure 2. (A) Mean reaction times (RTs) and (B) error scores (bottom panel) in Experiments 1 to 4, depicted separately for the different types of conjunction cues, as a function of whether cue and target were in the same location (valid trial) or different locations (invalid trial). As shown by the significant validity effect (i.e., faster RTs and fewer errors on valid than invalid trials), only conjunction cues that matched the relative features of the target attracted attention. * $p < .05$. ** $p \leq .01$. *** $p \leq .001$.

context cues matched the nontarget features (see Figure 2A, B). The two feature-match conditions produced inverse validity effects, with slower RT on valid than invalid trials (RelCol–: −31 ms, $t[13] = 3.07$, $p = .009$; RelSize–: –27 ms, $t[13] = 2.75$, $p = .016$), whereas cue validity did not affect performance in the nontarget-match condition (–14 ms), $t(13) = 1.52$, $p = .15$. With this, the results of Experiment 1 show that selection of a color-size conjunction cue depended on whether the cue matched the relative features of the target (e.g., bluest largest), not its feature values (e.g., medium, bluish green), in line with the relational account.

Experiment 2. Similar results were obtained in Experiment 2, in which the equiluminant colors of Experiment 1 were replaced with grayscale stimuli that varied in luminance. The same 7×2 ANOVA computed over the mean RTs showed significant main effects of the cue condition, $F(6, 72) = 9.48$, $p < .001$, $\eta^2 = .44$, validity, $F(1, 12) = 19.78$, $p = .001$, $\eta^2 = .73$, and a significant Cue Condition × Validity interaction, $F(6, 72) = 13.19$, $p < .001$, $\eta^2 = .52$. As shown in Figures 2A (blue line graph), cues in the full-match and relative-match conditions all attracted attention, as evidenced by significant validity effects (full match: $t[12] = 5.05$, $p < .001$; color match: $t[12] = 4.53$, $p = .001$; size match: $t[12] = 5.87$, $p < .001$; no match: $t[12] = 2.96$, $p = .012$). In turn, conjunction cues that were physically identical to the target but mismatched its relative features (feature-match cues) failed to attract attention and showed trends for inverse validity effects, with longer RT on valid than invalid trials. The inverse validity effect was also significant in the RelSize– condition, $t(12) = 3.09$, $p = .009$, but not in the RelCol– condition, $t < 1$. Similarly, the nontarget-match condition produced a trend for an inverse validity effect, which, however, failed to reach significance, $t(12) = 1.86$, $p = .087$.

Discussion

The first two experiments yielded several important findings. First, the results showed that an irrelevant conjunction cue can reliably capture attention in the spatial cueing paradigm. Validity effects were in fact similar in magnitude to those observed with salient cues in previous studies (67 and 62 ms for the fully matching cues in Experiments 1 and 2, respectively; Becker et al., 2013), which shows that attentional capture is not limited to salient cues, and that it does not depend on the cue having a unique elementary feature. Contrary to theories that reserve an important role for bottom-up saliency in attention (e.g., Theeuwes, 2010), these results demonstrate that top-down tuning to the target features alone is sufficient to elicit attentional capture by task-irrelevant cues.

Second, the present results provide the first evidence against the classical feature-based accounts of conjunction search and for a relational account of conjunction search. Contrary to feature similarity theories (e.g., Duncan & Humphreys, 1989; Martinez-Trujillo & Treue, 2004) and object-based accounts (e.g., Wyble et al., 2013), cues that perfectly matched the target features failed to attract attention (see feature-match conditions), whereas cues with target-dissimilar features reliably attracted attention (see relative-match conditions).

These results also cannot be attributed to the specific features of the conjunction cue: As half of the participants in each experiment

searched for a bluer target (or darker target) and the other half for a greener target (or lighter target), the critical cue (e.g., no-match cue) was blue (or dark) for half of the participants and green (or light) for the other half of participants, ruling out that the effects were mediated by perceptual factors.

Similarly, it cannot be argued that capture by the target-dissimilar cues (e.g., no-match cue) was due to the fact that they could not be distinguished from the target-similar cues. A failure to distinguish between target-similar and dissimilar colors (e.g., aqua and blue) would have resulted in no capture by any of the cues, as the cue displays contained both colors. In addition, the cues in the feature-match conditions often showed inverse cueing effects (see Figure 2A, B), with faster RT on invalid trials than valid trials. The inverse cueing effect could be due to the fact that the two context cues that matched the relative features of the target (e.g., bluest/largest) attracted attention (e.g., Harris et al., 2013), but other explanations are conceivable as well (e.g., Carmel & Lamy, 2014). More importantly, the inverse cueing effects indicate that target-similar and target-dissimilar features were readily distinguished.

Together, these findings demonstrate that the physical similarity between cue and target (or the cue context and the nontargets) is neither necessary nor sufficient for attentional capture, contrary to current feature-based and object-based accounts. Instead, capture by a conjunction cue is apparently independent of the physical similarity between cue and target and depends only on a match in the relative features of cue and target. With this, the present results are in line with the relational account, but not feature similarity accounts, object-based accounts, or saliency-based accounts of capture.

However, the results do not yet provide decisive evidence for a relational account of conjunction search, as they are still consistent with other feature-based accounts—most notably, an optimal tuning account (e.g., Navalpakkam & Itti, 2007). According to optimal tuning accounts, attention can be biased to an exaggerated target feature that is shifted away from the nontarget features (e.g., blue when the target is aqua), to increase the signal-to-noise ratio for the target signal (Navalpakkam & Itti, 2007; Scolari & Serences, 2010; but see Schoups, Vogels, Qian, & Orban, 2001). Optimal tuning could potentially account for the results of Experiments 1 and 2, because the conjunction cues in the relative-match conditions had feature values that were shifted away from the nontarget features (e.g., were larger/bluer than the target itself). Similarly, the target-similar cue in the feature-match conditions had to compete against cues with exaggerated target features, which could explain why they failed to capture attention.

The optimal tuning account is similar to the relational account in that both assume that the context of the target can determine how attention is tuned to the target. Deviating from the relational account, however, optimal tuning does not assume that feature relationships are computed or that they drive visual selection. Rather, the relatively matching cues captured attention because they were most similar to an exaggerated target feature value. In this respect, optimal tuning is still a feature-specific account. Experiment 3 was designed to test the optimal tuning account against the relational account, to determine whether conjunction search is truly relational or still feature-specific.

Experiment 3

The aim of Experiment 3 was to assess whether conjunction search is based on purely context-dependent selection mechanisms, by critically testing the relational account against an optimal tuning account. To that aim, Experiment 3 assessed whether conjunction cues that have the features of the *nontargets* can attract attention. According to optimal tuning accounts, a conjunction cue should capture attention once it matches the (exaggerated) feature values of the target, whereas cues that match the nontarget features should fail to attract attention. By contrast, the relational account maintains that cues that match the nontarget features should still be able to attract attention, provided that the conjunction cue has the same relative features to the other cues in the cueing display as the target has to the nontargets (e.g., bluest/largest).

In Experiment 3, the conjunction cues in the color-match and size-match conditions were identical to one set of nontarget items, and the cue in the no-match condition was a conjunction of two nontarget feature values (see Figure 1). Across all these conditions, the cue still had the same relative features as the target (bluest/largest), and thus the relational account would predict that the cue should still capture attention. By contrast, according to the optimal tuning account (and all other feature-specific accounts of attention), the cues in the relative-match conditions should fail to attract attention.

Method

Data. Sixteen new participants participated in Experiment 3. Two participants were excluded because of high error rates (>40%).

Stimuli, design, and procedure. The stimuli and procedures used in Experiment 3 were identical to the bluer search condition of Experiment 1, except that the conjunction cues in the relative-match conditions were now rendered similar or identical to the nontarget features (see Table 1 for a full description of the stimuli). The S, M, L, and XL bars in Experiment 3 measured $1.0° \times 0.2°$, $1.2° \times 0.3°$, $1.7° \times 0.4°$ and $2.2° \times 0.6°$, and in addition to the colors used in Experiment 1, the cue context could contain two "extragreen" cues (*xy*: .247, .451).

Results

As shown in Figure 2 (green line graph), rendering the conjunction cues more similar or identical to the nontargets did not alter the results pattern: The conjunction cues matching the relative features of the target still all attracted attention, regardless of whether they physically matched the target or the nontarget features. This was reflected in significant validity effects for the conjunction cue in the full-match condition, $t(13) = 9.34$, $p < .00$, and the three relative-match conditions (color match: $t[13] = 5.72$, $p < .001$; size match: $t[13] = 6.57$, $p < .001$; no match: $t[13] = 4.80$, $p < .001$). By contrast, conjunction cues that physically matched the target features, but not its relative features, all failed to attract attention (see feature-match conditions; Figure 2A). The cue in the nontarget-match condition similarly failed to attract attention and showed a significant inverse validity effect, $t(13) = 3.91$, $p = .002$.

Table 1

Overview of the Cue-Target Conditions Used in Experiments 3 and 4

Experiment 3: Cue varied.

Target: L2. // Nontargets: M2 + L3. (Target bluest/largest item in target display)

Cond	Full Match	Color Match	Size Match	No Match	Full Match/ RelCol-	Full Match/ RelSize-	Nontarget Match
Cue/ Cue Context	L2, among M2 + L3	M2, among S2 + M3	L3, among M3 + L4	M3, among S3 + M4	L2, among M2 + L1	L2, among XL2 + L3	M3, among M2 + L3

Experiment 4: Target / Nontargets varied (blocked).

Cue: M2. // Cue Context: S2 + M1. (Cue brightest/largest item in cue display)

Cond	Full Match	Color Match	Size Match	No Match	Full Match/ RelCol-	Full Match/ RelSize-
Target/ Nont.	M2, among S2 + M1	L2, among M2 + L1	M3, among S3 + M2	L3, among M3 + L2	M2, among S2 + M3	M2, among L2 + M1

Note. Conjunctions are specified by their combination of Size (S, M, L, XL) and Color or Luminance (Values: 1–4; see Legend of Fig. 1b). In Experiment 3, colors ranged from blue to green, and the size varied from S to XL, whereas in Experiment 4, stimuli varied in luminance (Values: 1–4). In Experiment 4, the cueing display was always constant, and only the size and luminance of the target and/or nontargets varied across different blocks, such that the cue matched the size and luminance of the target, a nontarget or neither (underscores indicate stimuli that were identical to the conjunction cue).

Discussion

The results from Experiment 3 shows that conjunction cues that match the relative features of the target attract attention even when they have the same features as the nontargets. These results argue against an optimal tuning account and show that the features of the conjunction cue are evaluated relative to the cue context. That is, whether a cue will capture attention is determined by whether its relative features match the target's relative features, not by the feature values of target and cue or their similarity.

These results are the first decisive evidence for a relational account of conjunction search. By implication, the results also show, for the first time, that it is possible to tune attention to two relative features in parallel (e.g., larger and bluer), and that it is possible to extract information about the relative features of the cues even when the display contains multiple stimuli with identical features, and the cue display is presented only briefly. Taken together, these results indicate that conjunction search is mediated by a rapid, relational selection mechanism that evaluates the relative features of the cues in parallel, and deploys attention to the cue that best matches the relative features of the target.

Experiment 4

In Experiment 4, we implemented the conditions of Experiment 3 and additionally recorded the EEG of participants to obtain more fine-grained information about the time course of attentional capture by conjunction cues. Specifically, we assessed the N2pc in response to the conjunction cue, which is an electrophysiological marker for attention defined as a higher contralateral than ipsilateral negativity over occipital/parietal electrodes (PO7, PO8) approximately 180 to 300 ms poststimulus onset (e.g., Eimer, 1996; Luck & Hillyard, 1994; Luck, Woodman, & Vogel, 2000; Woodman & Luck, 1999).

Our reasons for measuring the N2pc were twofold: First, studies using only behavioral measures such as RT and errors have often been criticized because behavioral effects could still be due to later processes concerned, for instance, with disengaging attention from a selected stimulus rather than the early orienting response (e.g., Schönhammer, Grubert, Kerzel, & Becker, 2016). The N2pc is widely accepted as one of the best indicators that attention was allocated to a stimulus. Thus, showing a significant N2pc in response to relatively matching conjunction cues would provide strong evidence that the conjunction cue attracts attention in virtue of its relative features, rather than only determining the dynamics of later, postselectional processes.

Second, measuring the N2pc provides detailed information about the time course of selecting conjunction stimuli, which allows evaluating a new two-stage account of conjunction search (Eimer & Grubert, 2014). Of note, Eimer and Grubert (2014) argued that selection of a conjunction stimulus proceeds in two stages; initially, feature-based attention is applied to all target-matching stimuli in the visual field, including stimuli that match the target only in one feature dimension (e.g., only color-matching or size-matching stimuli). Selection of a stimulus in virtue of its combined features occurs only later, by an object-based selection mechanism that operates on an integrated representation of the features of the conjunction cue. According to this two-stage account, the N2pc is composed of an early, feature-based component that is already present approximately 200 to 250 ms after display onset, and a later, object-based component that emerges 250 to 300 ms after display onset.

In the present study, selection of the conjunction cue could not be due to an early, feature-based mechanism, because the features of the conjunction cue were evenly distributed across all cues in the display (i.e., all items partially matched the conjunction cue), and therefore none of the cues could have been selected (as selection of all stimuli results in zero selection). Thus, the validity effect of the conjunction cue would have to be attributed to the later, object-based selection mechanism, which operates on an integrated representation of the target feature. This hypothesis, however, seems difficult to reconcile with the fact that conjunction cues captured attention in virtue of matching the relative features of the target. The relatively matching cues clearly constituted different objects as the target object (e.g., a blue, large item vs. an aqua, medium item), and, in Experiment 3, the cues had exactly the same features as the nontargets, which should have prevented capture by these cues.

The results of the present experiments seem more consistent with the framework of classical feature-based theories such as the *guided search model* (Wolfe, 1994). As mentioned above, guided search assumes that information from multiple dimensions is always combined at the level of the saliency map, across all search tasks (i.e., pop-out search, feature search and conjunction search; Wolfe, 1994). Hence, there is no reason to expect special delays in selecting conjunction stimuli (at least not when the conjunction stimulus attracts attention), and therefore the conjunction cue should elicit an N2pc within the same time window as observed for salient cues that have a unique feature value (e.g., 190 to 270 ms).

Because none of the previous studies have examined the N2pc to conjunction stimuli in the classical paradigm, in which the other stimuli in the context all share a feature with the critical conjunction stimulus, it is still an open question whether the N2pc to the conjunction stimulus will be within the standard range of feature search or delayed. In Experiment 4, we addressed this question by testing whether a significant N2pc to the conjunction cue could be observed in the same time window as observed for conjunctively defined stimuli that have a unique feature value, namely, 190 to 270 ms (Kiss et al., 2013).

To ensure that differences in the N2pc could not be attributed to any differences in the stimulus characteristics between the conjunction cues (e.g., Eimer, 1996), we used the same cuing display across all conditions. Six different conditions that corresponded to the first six conditions of the previous experiments were created by varying the target and nontarget colors across different blocks. As in Experiment 3, the conjunction cues in the size-match and color-match conditions were identical to one set of nontarget items (i.e., *not* shifted away from the nontarget feature values), and similar to Experiment 2, the conjunction stimuli were grayscale stimuli that varied in luminance (see Table 1 for a full description of the target–nontarget conditions).

If the behavioral validity effects for relatively matching conjunction cues in Experiments 1 to 3 were due to attentional capture (not later, postperceptual processes), then all four relative-match cues should elicit a significant N2pc, whereas the two feature-match cues should fail to elicit a significant N2pc. Moreover, if attention is guided to the conjunction cue via an early, feature-based process, the N2pc to the relatively matching cues should be observed within the standard time window (190 to 270 ms postcue

onset). By contrast, if selection of a conjunction cue is mediated by object-based processes, then only the perfectly matching full-match cues should elicit a significant N2pc, and the N2pc could also be significantly delayed (e.g., 250 to 300 ms poststimulus onset; Eimer & Grubert, 2014).

Method

Data. Data were collected from 14 new, right-handed subjects, who were reimbursed with $20.00 for their participation. One participant was excluded because of high alpha activity. Error trials were excluded in the mean RT but included in the N2pc results, because (a) the N2pc was measured in response to the cues and should remain unaffected by the manual response to the target, and (b) errors varied significantly across conditions, so that excluding error trials would have created noise in a subset of conditions (see Figure 2). After exclusion, on average, 110 trials were available for the computation of ipsi- and contralateral waveforms, respectively (range = 65 to 120; average loss of data = 10.2%).

Apparatus. The continuous EEG was recorded using a 64-channel ActiveTwo Biosemi EEG system (Biosemi Instrumentations, Amsterdam, the Netherlands), digitized at 1,024 Hz. All electrodes were mounted in an elastic cap corresponding to the 10–10 system, and impedances were kept below 5kΩ. EEG data were referenced to the common average and filtered at 45 Hz. Trials with artifacts (eye movements, or scalp channels exceeding ± 75 μV) were excluded from the EEG data analysis (see Becker et al., 2014, for a similar procedure).

Stimuli, design, and procedure. All cue displays consisted of a medium-sized gray conjunction cue presented among two small gray cues and two medium dark gray cues. Six different cue-target conditions were created by varying the target and nontarget features across six different blocks, with the block order randomized across participants (see Table 1). The S, M, L bars measured $1.3° \times 0.3°$, $2.3° \times 0.5°$ and $3.0° \times 0.8°$, respectively, and the dark gray, medium gray, and light gray bars had a luminance of 17.2 cd/m², 57.1 cd/m², and 77.4 cd/m², respectively. To prevent large event-related potentials (ERPs) in response to the onset of the placeholders, only a portion of the lower line of the placeholder boxes (0.4°) was presented. To optimize conditions for the measurement of cue-related ERPs (Eimer, 1996), the duration of the fixation display was increased to 200 ms, which increased the cue-target interval to 300 ms (100 ms cue display, 200 ms fixation display, 100 ms target display). Moreover, to maximize the number of relevant trials for the analysis of the N2pc, the conjunction cue was always presented on the left or right side of the display, and the number of trials was increased to 240 trials per block (1,440 trials in total). The conjunction cue was always nonpredictive of the target location, yielding 40 valid trials and 200 invalid trials per condition.

Results

Mean RTs. The behavioral results of Experiment 4 are depicted in Figure 2 (gray line graph). The results of a 6×2 ANOVA computed over the mean correct RT revealed significant main effects of the cue-target condition, $F(5, 60) = 4.61$, $p = .004$, $\eta^2 = .27$, cue validity, $F(1, 12) = 16.76$, $p = .001$, $\eta^2 = .58$, as well as a significant interaction between the variables, $F(5, 60) = 7.18$, $p = .001$, $\eta^2 = .37$. Paired t tests revealed significant validity effects in the full-match condition, $t(12) = 7.34$, $p < .001$, and the three relative-match conditions (color match: $t[12] = 5.10$, $p < .001$; size match: $t[12] = 3.04$, $p = .010$; no match: $t[12] = 2.80$, $p = .016$). No significant validity effects were observed in any of the feature-match conditions, in which the conjunction cue had the same physical features as the target but failed to match one of its relative features (all $ps > .22$; see Figure 2A, B).

N2pc. The mean difference wave and the individual EEG waveforms for the conjunction cues in the six cue conditions are depicted in Figure 3. For the analysis of EEG data, we computed the mean amplitudes of the N2pc (i.e., contralateral minus ipsilateral negativity at electrodes PO7, PO8), relative to a prestimulus baseline of -100 ms to 0 ms before cue onset, in the "standard" time window (190 to 270 ms postcue onset). A one-way ANOVA revealed significant differences between the cue-related N2pc amplitudes across the six cue conditions, $F(5, 60) = 5.6$, $p < .001$. Pairwise comparisons showed that there was a significant N2pc for the cues in the full-match condition, $t(12) = 4.45$, $p \leq .001$, the color-match condition, $t(12) = 3.89$, $p = .002$, and the no-match condition, $t(12) = 2.23$, $p = .045$, whereas the N2pc for the size-match condition was only marginally significantly different from zero, $t(12) = 2.14$, $p = .054$. The cues in the two feature-match conditions failed to elicit a significant N2pc, and instead both had *positive* mean amplitudes, which, however, did not differ significantly from zero, both $ps > .09$. These results provide independent evidence for the view that capture by a conjunction cues depends on whether it matches the relative, not physical, features of the target.

To test whether the time course of attending to the cue differed across conditions, we also compared the mean N2pc onset latencies (measured as the duration from the onset of the cue display to 50% of the individual N2pc peaks). However, a one-way ANOVA of the jack-knifed onset latencies (Ulrich & Miller, 2001) failed to show any significant differences between N2pc latencies, $F_{corr} < 1$. The same result was obtained when the analysis was limited to the first four conditions that showed a significant N2pc, $F_{corr} < 1$, and when the jack-knifed N2pc peak latencies were analyzed, $F_{corr} = 1.1$, $p = .36$.

Discussion

The results of Experiment 4 support a relational account of conjunction search, by showing that selection operates on the relative, rather than the physical, features of a conjunction stimulus: Behavioral cueing effects and a significant N2pc was observed only for conjunctions that matched the relative features of the target, not for cues that were identical to the target but failed to match its relative features. As in Experiment 3, the relatively matching conjunction cues had the same features as the nontargets, and thus should not have attracted attention according to current feature-specific accounts of attention (including the optimal tuning account). Moreover, the results cannot be attributed to any perceptual differences between the cue conditions, as the cue display was the same across all conditions and only the target displays varied across different blocks. With this, the observed relational effects can be clearly attributed to top-down controlled processes.
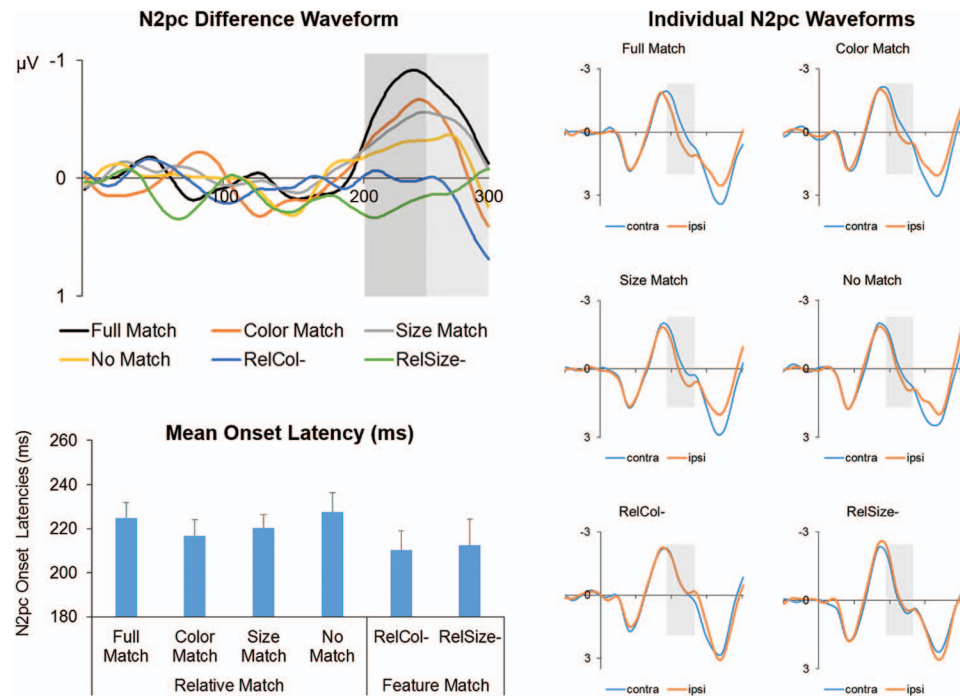
Figure 3. Left top: ERP Difference waves (contra – ipsilateral) in response to the conjunction cue in the time window 0 to 300 ms, depicted separately for the six conditions of Experiment 4. The dark shaded area indicates the early, 200- to 250-ms time window, and the light shaded area a later, 250- to 300-ms time window. Left bottom: Mean N2pc onset latencies (at 50% of the N2pc peak) in the six conditions (with error bars indicating the standard error of the mean). Right panels: Contra- and ipsilateral waveforms in response to the conjunction cue depicted from −100 to +400 ms, relative to cue onset. For the figures, data were averaged over a 25-ms time window with a moving average filter to reduce noise, whereas the analyses were all based on unfiltered data.

Of note, the N2pcs were numerically larger in the full-match condition, in which the cues matched both the relative and physical features of the target, than in the relative-match conditions, in which the cues matched only the relative features of the target (color match, size match, no match). These differences cannot be attributed to characteristics of the N2pc, as exactly the same trends were observed in the behavioral validity effects (see Figure 2). Stronger capture by target-similar cues may indicate that (some) observers may have adopted a more feature-specific search strategy in a subset of the conditions (e.g., toward the end of the experiment), possibly because the target's relative features switched across blocks (i.e., between larger/smaller, and lighter/darker; see Table 1 and Harris et al. [2013], for related results). Another possibility is that, due to the longer cue-target delay in Experiment 4 (200 ms), the time course of disengaging attention from the selected cue played a larger role in the behavioral and electrophysiological results. Previous visual search studies have shown that perfectly matching distractors hold the gaze for longer than relationally matching but target-dissimilar distractors, possibly because distractors with deviating colors can be quickly rejected (Becker et al., 2014). Thus, perfectly matching cues may have produced a larger validity effect and N2pc, because attention dwells longer at the location of perfectly matching cues. According to this account, the initial orienting response would be determined by a match in the relative features, with later, feature-based pro-

cesses contributing mainly afterward to validity effects and the N2pc. This account could explain why a perfectly matching cue modulated the N2pc when it matched the relative properties of the target, but completely failed to elicit an N2pc when it did not match the relative properties of the target (see Figure 2). Although the exact cause for this possible feature-specific modulation of capture cannot be ascertained with certainty, it is important to note that the results still showed clear evidence for relational search dominating the results, as capture by the conjunction cues was observed if, and only if, the relative color and size of the cue matched the relative features of the target.

A third important finding relates to the time course of capture by the conjunction cues. Contrary to the two stage account of Eimer and Grubert (2014), there was no evidence that selection of conjunction stimuli is initially determined by feature-based processes and later by object-based processes. According to this view, capture by relatively matching cues should have been limited to an early time window (200 to 250 ms; see dark shaded area of Figure 3), with a later time window showing strong evidence for capture by perfectly matching cues (250 to 300 ms; light shaded area in Figure 3). However, the results did not show any evidence for corresponding differences in the time course of the N2pc across the different conditions. Instead, the N2pcs had a similar time course as that observed for simple feature and pop-out stimuli (e.g., Kiss et al., 2013). These results are more consistent with current feature-based theories such as guided search (e.g., Lee et al.,

1999; Wolfe, 1994), which propose that conjunction search relies on the same processes that also govern pop-out search and feature search.

General Discussion

The present study yielded several important findings. First, the results show that an irrelevant conjunction cue can attract attention in a variant of the Posner cueing paradigm (Posner, 1980), even when it is presented only briefly, and does not have any unique elementary features or any usable feature contrast. In previous studies using the spatial cueing paradigm, the conjunction cue was typically salient (e.g., Kiss et al., 2013), and in visual search tasks, displays were more sparsely populated, and/or the target still had a unique feature value on a subset of trials (e.g., Eimer & Grubert, 2014). By contrast, in the present study, both the conjunction target and the cue were genuine feature conjunctions that differed *only* in their combination of features from other stimuli in the surround. Hence, the finding that the irrelevant cue still reliably attracted attention shows that top-down tuning of attention is sufficient to elicit reflexive orienting, and argues against accounts that explain attentional capture with respect to bottom-up, stimulus-driven processes (e.g., Li, 2002; Theeuwes, 2010). With this, the present study extends on the previous evidence that conjunction search can be efficient (e.g., He & Nakayama, 1992), by showing that genuine conjunction stimuli can also reflexively attract attention (provided that they match the discriminating attributes of the conjunction target).

A second important finding of the present study is that attentional capture by conjunction cues was governed by tuning to the relative features of the conjunction target. Contrary to prior formulations of the relational account, the target evidently does not have to differ in a single relative feature from all other items (i.e., bluest or largest) for relational search to occur (Becker, 2010; Brand et al., 2014). This indicates the need to reformulate the limitations of relational search. Specifically, the present results indicate that relational search will be observed once (a) the relative feature(s) of the target are reliable (i.e., the target does not vary between bluer/greener or larger/smaller across trials), and (b) the target has *one of* the relatively extreme features in the display (i.e., does not differ in two opposite relative features from the nontargets within a trial). Both of these conditions are met in the most frequently used conjunction search tasks (e.g., when a red vertical target is presented among red horizontal and green vertical nontargets), indicating that in these previous studies, conjunction search may have been accomplished by tuning to the relative features of the target, not to its specific features. In line with this possibility, the current data show that observers prefer relational search even when it is equally possible to tune attention to the feature values of the target. The same preference for relational search has been observed in previous studies in search for pop-out targets and targets with a unique feature, across a wide range of stimulus conditions (i.e., search for size, shape, luminance and color; e.g., Becker, 2010, 2013; Becker et al., 2013). The present study shows that relational search is also preferred in the ecologically more valid conditions of conjunction search, which indicates that the visual system may be more likely to operate in a context-dependent rather than feature-specific search mode in natural environments.

In addition, the present results indicate that feature relationships have an important dual role in selection. First, the relative features of the target determine how attention is tuned to the target (e.g., larger/smaller). This kind of context-dependent tuning of attention

is also proposed in the optimal tuning account (Lee et al., 1999; Navalpakkam & Itti, 2007; Scolari & Serences, 2010). Extending on the optimal tuning account, we showed that reflexive orienting to irrelevant stimuli depends only on *their relative features*, and is independent on whether they match or mismatch the physical feature values of the target. This conclusion is supported by two observations, namely, (a) that a cue with the same features as the nontargets still captured attention (see Experiments 3 and 4), and (b) that a target-dissimilar cue captured attention even when it was presented among other items that had exactly the same feature values as the target (see Experiments 1 and 2). Thus, the relative attributes of the irrelevant cue determined whether the cue was attended or not. These results are inconsistent with current feature-specific accounts of conjunction search, and show that information about the relative features of all items is extracted rapidly and automatically (i.e., when the stimuli are presented only briefly, and when they are irrelevant to the task), in line with a relational account of conjunction search. The present findings and the results of Brand et al. (2014) show that relative features can still be extracted even when none of the items in the display has a unique feature, and when they differ on irrelevant stimulus dimensions (e.g., the relative size of items can be used for guidance even when the stimuli also vary in color and orientation; Brand et al., 2014). These findings provide very strong evidence for the existence of a mechanism that rapidly and automatically extracts information about feature relationships.

A third important finding was that irrelevant conjunction stimuli were selected at an early point in time, within the range typically observed for more salient stimuli (e.g., Töllner, Müller, & Zehetleitner, 2012). Contrary to the dual stage account (Eimer & Grubert, 2014), there was no evidence that the N2pc for conjunction stimuli was delayed, and thus no support for the view that selection of a conjunction is based on later, object-based processes of attention. In this respect, it is important to notice that these findings are not inconsistent with previous results, as studies comparing feature versus conjunction search have not reported a significant delay in the N2pc in conjunction search (e.g., Eimer & Grubert, 2014; Luck & Ford, 1998; Luck, Girelli, McDermott, & Ford, 1997). Rather, the dual-stage account of Eimer and Grubert (2014) was based on the observation that partially matching nontargets (that match only one feature of the target) can also elicit a significant N2pc. Specifically, when the N2pc of the partially matching distractors is subtracted from the N2pc of a fully matching target stimulus, the remaining N2pc to the target is only significantly different from zero in a later time window (250 to 300 ms). This finding was taken as evidence that at an early stage (150 to 200 ms), attention is dominated by feature-based processes that indiscriminately select all partially matching stimuli, and that selection of a conjunction stimulus (at the exclusion of selecting other objects) requires object-based attentional processes that operate at a later stage of processing (Eimer & Grubert, 2014).

In the present study, early feature-based selection would have resulted in selection of all locations, and thus no significant N2pc, as all locations included a partially matching cue. Yet we found a significant N2pc to the conjunction cue within an early time window (150 to 200 ms), and thus significantly earlier than predicted by this subtraction model. Although this result seems at odds with the dual stage account, it does not contradict Eimer and Grubert's (2014) findings, as they did not test whether attention to

a conjunction stimulus would be delayed when it is in direct competition with partially matching distractors.

One possibility is that the subtraction logic, even though it has been extensively used in research (e.g., Chelazzi, Miller, Duncan, & Desimone, 1993; Luck et al., 1997), does not reliably model the underlying neural processes and thus, does not accurately predict the time course of the N2pc in conditions of direct competition. A second possibility is that attention to partially matching stimuli in the study of Eimer and Grubert (2014) was due to observers tuning attention to relative features. Specifically, the reduced N2pc to partially matching distractors in Eimer and Grubert could have been due to the fact that some of these distractors matched the relative features of the target (leading to a full-sized N2pc), whereas others failed to match one of the relative features of the target (leading to a zero N2pc). Of note, in the study of Eimer and Grubert (2014), three of the four colors and shapes were quite similar to each other and varied gradually along a continuum (i.e., blue, cyan, green varied along a blue-green continuum, and the shapes varied in the number of corners). Previous studies have shown a strong preference for relational search when stimulus features vary gradually and are clearly related (e.g., Becker, 2013; Becker et al., 2013). Hence, the reduced N2pc to partially matching stimuli could have resulted from mixture of capture and no-capture trials, thus explaining the reduced capture effect of partially matching distractors in Eimer and Grubert.

An important question is whether feature relationships or feature values constitute the most basic unit of information, and whether information about relative features can be extracted from feature-specific information, or vice versa. However, the present study seems to be consistent with both views. Most current theories of attention postulate that selection is based on sensory neurons that respond rather narrowly to a small range of specific feature values (*feature maps* or *feature detectors*; e.g., Treisman & Gelade, 1980; see also Lee et al., 1999; Navalpakkam & Itti, 2007). It is likely that information about feature relationships could be extracted from such feature-specific neurons in secondary processing steps (e.g., Becker, 2014; Meeter & Olivers, 2014). For instance, the results of Experiments 3 and 4 could be explained by assuming that attention is tuned very broadly to all feature detectors that respond more strongly to the target feature than the nontarget feature (not only to target-similar features or an optimal feature value). As this can include feature detectors that still respond sufficiently strongly to the nontarget feature values to elicit attention shifts (when tuning curves are broad), it would explain how nontarget-similar cues can attract attention (see Becker, 2014; Meeter & Olivers, 2014, for more detailed explanations). Still, the reverse also seems to be true, that information about the relative features of items could be used to extract feature-specific information in a secondary processing step (see Becker, 2014, for details). The mounting evidence that relational search is preferred to feature-specific tuning in a broad range of search tasks, and that information about relative features is extracted automatically and in parallel by early visual processes indicates that feature detectors may in fact be "relational feature detectors" that signal the presence of relative features rather than absolute feature matches. However, clearly, further research is needed to find out whether feature-specific information is extracted from broader, relative coding, or vice versa.

In conclusion, the present study found that attention can be tuned to the relative attributes of conjunction stimuli, so that attentional capture by irrelevant conjunction stimuli is not determined by target similarity, but by the context-dependent features of cue and target. In addition, we established that an irrelevant stimulus does not need to have a unique feature or a high feature contrast to attract attention. The view that conjunction stimuli are selected in virtue of an (early or late) object-based attentional mechanism could not be supported either. Instead, attentional capture by the conjunction cue depended only on whether it matched the relative features of the conjunction target (both features). Taken together, these results indicate a selection mechanism that rapidly computes feature relationships across all stimulus locations (cue and target locations), with signals from relatively matching cues being combined across different stimulus dimensions (i.e., of color and size). With this, it is likely that relational search is also applied in natural environments.

References

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

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Distractor heterogeneity versus linear separability in colour visual search. *Perception, 25,* 1281–1293. http://dx.doi.org/10.1068/p251281

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,* 247–265. http://dx.doi.org/10.1037/a0018808

Becker, S. I. (2013). Simply shapely: Relative, not absolute shapes are primed in pop-out search. *Attention, Perception, & Psychophysics, 75,* 845–861. http://dx.doi.org/10.3758/s13414-013-0433-1

Becker, S. I. (2014). Guidance of attention by feature relationships: The end of the road for feature map theories? In M. Horsley, M. Eliot, B. A. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research* (pp. 37–49). Switzerland: Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-02868-2_3

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. http://dx.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,* 634–647. http://dx.doi.org/10.1177/0956797612458528

Becker, S. I., Grubert, A., & Dux, P. E. (2014). Distinct neural networks for target feature versus dimension changes in visual search, as revealed by EEG and fMRI. *NeuroImage, 102,* 798–808. http://dx.doi.org/10.1016/j.neuroimage.2014.08.058

Becker, S. I., Harris, A. M., Venini, D., & Retell, J. D. (2014). Visual search for color and shape: When is the gaze guided by feature relationships, when by feature values? *Journal of Experimental Psychology: Human Perception and Performance, 40,* 264–291. http://dx.doi.org/10.1037/a0033489

Brand, J., Oriet, C., Johnson, A. P., & Wolfe, J. M. (2014). Flexible cue combination in the guidance of attention in visual search. *Acta Psychologica, 153,* 129–138. http://dx.doi.org/10.1016/j.actpsy.2014.10.002

Carmel, T., & Lamy, D. (2014). The same-location cost is unrelated to attentional settings: An object-updating account. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1465–1478. http://dx.doi.org/10.1037/a0036383

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature, 363,* 345–347. http://dx.doi.org/10.1038/363345a0

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45,* 891–900. http://dx.doi.org/10.1016/j.visres.2004.10.004

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501–517. http://dx.doi.org/10.1037/0096-3445.113.4.501

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. http://dx.doi.org/10.1037/0033-295X.96.3.433

D'Zmura, M. (1991). Color in visual search. *Vision Research, 31,* 951–966. http://dx.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. http://dx.doi.org/10.1037/0096-1523.10.1.32

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography & Clinical Neurophysiology, 99,* 225–234. http://dx.doi.org/10.1016/0013-4694(96)95711-9

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. http://dx.doi.org/10.1037/a0037387

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,* 847–858. http://dx.doi.org/10.1037/0096-1523.24.3.847

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,* 1030–1044. http://dx.doi.org/10.1037/0096-1523.18.4.1030

Harris, A. M., Remington, R. W., & Becker, S. I. (2013). Feature specificity in attentional capture by size and color. *Journal of Vision, 13,* 12. http://dx.doi.org/10.1167/13.3.12

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359,* 231–233. http://dx.doi.org/10.1038/359231a0

Hodsoll, J., & Humphreys, G. W. (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. http://dx.doi.org/10.3758/BF03194447

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489–1506. http://dx.doi.org/10.1016/S0042-6989(99)00163-7

Kahneman, D., & Treismann, A. (1984). Changing views on attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). Orlando, FL: Academic Press.

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. http://dx.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. http://dx.doi.org/10.3758/s13414-012-0391-z

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience, 2,* 375–381. http://dx.doi.org/10.1038/7286

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6,* 9–16. http://dx.doi.org/10.1016/S1364-6613(00)01817-9

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences, 95,* 825–830.

Luck, S. J., Girelli, M., McDermott, M. T., & Ford, M. A. (1997). Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cognitive Psychology, 33,* 64–87. http://dx.doi.org/10.1006/cogp.1997.0660

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31,* 291–308. http://dx.doi.org/10.1111/j.1469-8986.1994.tb02218.x

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4,* 432–440. http://dx.doi.org/10.1016/S1364-6613(00)01545-X

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,* 902–912. http://dx.doi.org/10.1037/0096-1523.28.4.902

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

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences, 29,* 317–322. http://dx.doi.org/10.1016/j.tins.2006.04.001

Meeter, M., & Olivers, C. N. (2014). Target features and target-distractor relation are both primed in visual search. *Attention, Perception, & Psychophysics, 76,* 682–694. http://dx.doi.org/10.3758/s13414-013-0611-1

Nako, R., Grubert, A., & Eimer, M. (2016). Category-based guidance of spatial attention during visual search for feature conjunctions. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 1571–1586. http://dx.doi.org/10.1037/xhp0000244

Navalpakkam, V., & Itti, L. (2006). An integrated model of top-down and bottom-up attention for optimal object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), 1–7.

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

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32,* 3–25. http://dx.doi.org/10.1080/00335558008248231

Schönhammer, J. G., Grubert, A., Kerzel, D., & Becker, S. I. (2016). Attentional guidance by relative features: Behavioral and electrophysiological evidence. *Psychophysiology, 53,* 1074–1083. http://dx.doi.org/10.1111/psyp.12645

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature, 412,* 549–553. http://dx.doi.org/10.1038/35087601

Scolari, M., & Serences, J. T. (2010). Basing perceptual decisions on the most informative sensory neurons. *Journal of Neurophysiology, 104,* 2266–2273. http://dx.doi.org/10.1152/jn.00273.2010

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135,* 77–99. http://dx.doi.org/10.1016/j.actpsy.2010.02.006

Töllner, T., Müller, H. J., & Zehetleitner, M. (2012). Top-down dimensional weight set determines the capture of visual attention: Evidence from the PCN component. *Cerebral Cortex, 22,* 1554–1563. http://dx.doi.org/10.1093/cercor/bhr231

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

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. http://dx.doi.org/10.1016/0010-0285(80)90005-5

Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology, 38,* 816–827. http://dx.doi.org/10.1111/1469-8986.3850816

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123,* 146–160. http://dx.doi.org/10.1037/0096-3445.123.2.146

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

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

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37,* 25–43. http://dx.doi.org/10.1016/S0042-6989(96)00111-3

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature, 400,* 867–869. http://dx.doi.org/10.1038/23698

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. http://dx.doi.org/10.1037/a0030517

Wyble, B., & Swan, G. (2015). Mapping the spatiotemporal dynamics of interference between two visual targets. *Attention, Perception, & Psychophysics, 77,* 2331–2343. http://dx.doi.org/10.3758/s13414-015-0938-x

---

Members of Underrepresented Groups:
Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.

- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.

- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, "social psychology" is not sufficient—you would need to specify "social cognition" or "attitude change" as well.

- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit http://www.apa.org/pubs/authors/review-manuscript-ce-video.aspx.