Australian
Psychologist

APS Australian Psychological Society

INVITED CONTRIBUTION

Why You Cannot Map Attention: A Relational Theory of Attention and Eye Movements

Stefanie I Becker

The University of Queensland

How do we select relevant information from cluttered visual environments? Current theories of attention and eye movement control have proposed a framework of feature detectors that allows us to attend to potentially important information and ignore irrelevant information. Each feature detector supposedly signals the existence of a single elementary feature (e.g., red) in the environment and can be modulated by the intention to search for particular features. In contrast to these feature detector accounts, recent visual search studies show that features are encoded relative to other features in the surround, in a context-dependent manner. Attention and eye movements are attracted to items with appropriate relative features (e.g., redder/larger/darker), regardless of their feature value. These findings question the feature detector concept and argue for a *relational account* of attention and eye movements.

Key words: attention; eye movement; gaze; relational theory; visual search.

Introspectively, we may have the impression that we can see and appreciate all objects in a visual scene at once. However, this impression is deceptive: Research has shown that we have to allocate attention to objects to become aware of them. Attention selects items for further in-depth processing, and is severely capacity limited. According to some theories, we can select and process only a single item at any given moment in time, which can then guide our behaviour (e.g., Huang, Treisman & Pashler, 2007). Our vivid impression of a rich visual scene is presumably due to a scanning process, during which attention is serially deployed to different locations in the visual field, and information about the objects at these locations is pieced together. Given that conscious perception of one object works at the exclusion of perceiving other—and potentially more important—objects, it is important to identify the factors that determine which items will be selected first. Correspondingly, much research has been devoted to examine the factors that guide visual attention and eye movements.

Previous research has shown that visual selection is determined by both stimulus-driven and goal-driven factors: Attention can be attracted to item with a high feature contrast in a stimulus-driven manner (e.g., Theeuwes, 1992), or it can be guided to task-relevant stimuli in a goal-driven manner. For instance, the intention to find an object such as a red Jeep can guide attention to locations in the visual field containing red items (e.g., Folk & Remington, 1998; Folk, Remington, & Johnston, 1992; Wolfe, 1994).

According to current theories of attention, top-down control over visual attention is achieved by modulating the response gain of so-called *feature detectors* or *feature maps*, which are topographically organised neurons that respond to specific feature values such as colours or orientations. Attention is guided by the summed signal from all feature detectors, whereby we have a dedicated feature detector (or topographic map) for all feature values that can be pre-attentively distinguished; including different orientations (e.g., 0°, 20°, 40°), colours (e.g., red, green, blue, yellow), and motion directions (e.g., Lee, Itti, Koch, & Braun, 1999; Treisman & Sato, 1990; Wolfe, 1994). The intention to search for a particular feature value can enhance the response gain (firing rate) of feature detectors responding to the target (*target activation*) and/or attenuates the response gain of feature detectors responding to irrelevant features (*non-target inhibition*). Target activation and non-target inhibition allow limiting selection to task-relevant features (see Figure 1).

The Relational Theory of Attention

Contrary to the feature detector theories of visual attention, I recently developed a *relational account*, which claims that attention and eye movements are determined by feature *relationships*. Feature relationships specify how the feature of a sought-after item differs from the features of the irrelevant context (e.g., redder, darker, larger; Becker, 2010). According to the relational account, the target-defining feature (e.g., orange) is not processed independently of the features of the non-targets, but relative to them, resulting in context-dependent processing of all features. Attention is tuned to the relative attributes of the target, and items attract attention in virtue of matching the feature relationships that the target has to the (majority of) non-target features (see Figure 2).

Correspondence: Stefanie I. Becker, School of Psychology, The University of Queensland, St Lucia, QLD 4072, Australia. email: s.becker@psy.uq.edu.au

This research was supported by the Australian Research Council (ARC) Discovery Grant and fellowship DP110100588, awarded to S.I.B.

Accepted for publication 30 April 2013

doi:10.1111/ap.12028

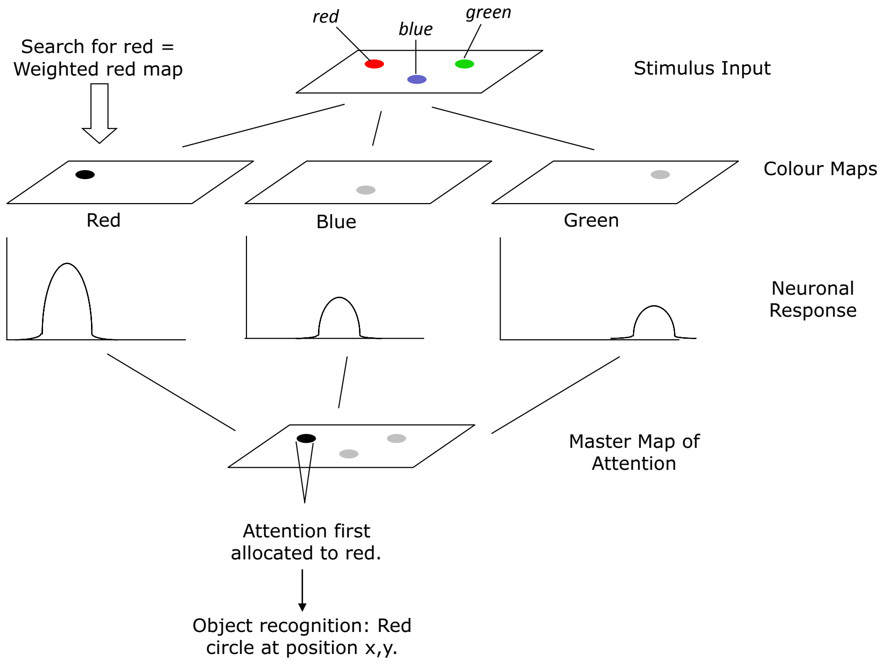

Figure 1 Schematic Description of Feature Map Theories. *Note*. In search for a red target, the response gain of the topographic red feature map (i.e., neurons responding to red) is enhanced. When information from all feature maps is integrated on the overall master map of attention, red stimuli have the highest activation value. Attention is first drawn to the location containing the highest activation. The example depicts only a subset of colour maps; similar feature maps have been proposed for different orientations, motion directions, curvature, and depth (e.g., Lee et al., 1999; Navalpakkam & Itti, 2006; Treisman & Gelade, 1980; Wolfe, 1994).

Figure 2 Example for Searching for Orange in Different Contexts. *Note.* According to the relational account, colours are encoded relative to the context, and visual selection is biased to the relative attribute that allows distinguishing the target from the non-targets. Hence, when searching for the orange-clad goalkeeper, visual selection would be biased to *redder*, or the *reddest* item, in the picture on the left, whereas attention would be biased to *yellower*, or the *yellowest* item in the picture on the right (see, e.g., Becker, 2010, 2013a,b; Becker et al., 2010, 2013).

The relational account centrally claims that attention is guided by relative attributes and not absolute feature values. In the relational vector account (Becker, 2010), features such as different colours are represented in a continuous feature space (e.g., CIE colour space), and the direction in which a given feature differs from other features is described by the direction of vectors pointing from each feature to the other features. When searching for a particular colour, attention would not be limited to those items that have exactly the same colour as the target, but to those items that have the same relative colour as the target (Becker, 2010). For example, in search for an orange item, attention would be biased towards *redder* items, or the *reddest* item, when the item is known or expected to occur in a context of yellow(er) items, whereas attention would be biased towards *yellower* or the *yellowest* item when the context is known or suspected to contain mostly red(der) items (see Figure 2).

One advantage of the relational account is that it does not require proposing a multitude of different feature detectors for each feature value that can be top-down selected, which has been recognised to be a major problem of feature detector theories. Of note, visual selection is quite fine-grained (e.g., Navalpakkam & Itti, 2006) and the feature detector theories require a large number of feature detectors (i.e., topographically organised sensory neurons)—probably more than can reasonably be expected to subserve the guidance of attention (cf. Maunsell & Treue, 2006). A second advantage of the relational account is that it seems to offer a more flexible and adaptive search criterion in the variable conditions of natural environments. An object with a specific size, colour or luminance value can look very different under different lighting conditions, viewing angles, and the observer's distance from the object, whereas the relative properties are often invariant to such changes (Becker, 2013a; Foster, 2003). This is especially true for objects of different sizes: Whereas an object's size greatly varies with the distance of the observer to the object, the relative size of an object to other objects (larger/smaller) is largely distance-invariant. Hence, it seems that encoding the relative colour, size or brightness of an object offers a more robust and versatile top-down tuning strategy in a variable and unpredictable environment than encoding the feature values of objects. Third, feature relationships could also help explain the learning process for gaining control over attention. To date, it is largely unknown how we learn to tune attention to the appropriate features; or in other words, it is unknown how the top-down connection from "intention to search for feature x" to "enhancing the response gain of neural population y" is established (e.g., Ahissar & Hochstein, 2004). However, given the large number of feature detectors, an explanation relying on connections between specific colour words and feature detectors could prove challenging because of the implied combinatorial explosion (e.g., Rumelhart, Hinton, & Williams, 1986). Feature relationships can be successfully applied in a multitude of different contexts and may thus provide a more fruitful approach to understanding how such connections are established during learning.

According to the relational account, attention does not have to be tuned to particular colour values: When we are trying to find an object of a particular shade of pink or aquamarine, the visual system could quickly evaluate how this colour should differ from a given context, to tune attention to the relative colour (e.g., redder/bluer or greener/bluer) that will allow distinguishing the colour from the other colours in the context. Naturally, such a "relational search strategy" would require the ability to quickly assess the attributes that are dominant in the context of a visual scene. Are we indeed able to make such quick judgements of the entire scene content? Previous research has shown that we are in fact able to make such quick assessments that are also surprisingly accurate (e.g., Baldassi & Burr, 2000; Chong & Treisman, 2005a,b; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). After a very brief look (e.g., 30ms) at an array with stimuli of varying colours, sizes or orientations, observers can very accurately report the average colour, size or orientation of all stimuli in the display (e.g., Chong & Treisman, 2003, 2005a,b; Jacoby, Kamke, & Mattingley, 2013). Accurate averaging judgements are moreover not limited to elementary features (e.g., orientation, motion, colour), but have also been reported for higher-order perceptual features such as the average emotional expression of a crowd (e.g., de Fockert & Wolfenstein, 2009; Haberman, Harp, & Whitney, 2009; Sweeney, Grabowecky, Paller & Suzuki, 2009). The ability to assess average scene contents is thought to rely on a summary statistical representation of a scene, which is available very early. In the relational account, the summary statistical representation could aid visual selection by defining the average properties of the context, which can then be used to tune attention to those feature relationships that will allow discriminating the target from the context. It is currently unknown whether this representation indeed plays a role in tuning attention to feature relationships. However, as described in more detail below, previous studies have demonstrated that feature relationships can guide attention to items differing in size, colour and shape (e.g., Becker, 2010, 2013a; Becker, Folk, & Remington, 2010, 2013).

Evidence for the Relational Account

Interestingly, feature relationships have been demonstrated to determine visual selection both in effects commonly attributed to "bottom-up" processes as well as effects presumably driven by "top-down" controlled processes. In the following, the evidence for the relational account will be described separately for bottom-up versus top-down processes.

"Bottom-up" Feature Priming Is Relational

Visual search is typically faster when the target from the previous trial is repeated than when it changes. For instance, in visual search for a pop-out target that was randomly either red or green, Maljkovic and Nakayama (1994) found that responses were faster when the colour of the target from the previous trial was repeated (e.g., red target preceded by red target) than when the target and non-target colours had switched (e.g., red target preceded by green target, among red non-targets). Corresponding intertrial effects have been found in search for other features, including shape, size, luminance, position and orientation search tasks, and are commonly regarded as bottom-up because advance information about the target cannot eliminate these intertrial effects (Becker, 2008a; Hillstrom, 2000; Leonard & Egeth, 2008; Maljkovic & Nakayama, 1994).

According to the *priming of pop-out hypothesis* (Maljkovic & Nakayama, 1994), attending to the target on a given trial *primes* or *biases* attention to select the same feature on subsequent trial(s). The selection bias for the previously selected target feature facilitates search when the target feature is repeated and impairs target selection when the target and non-target features switch, compared to the previous trial ("switch costs").

In line with this view, several eye movement studies found that switching the target and non-target features produces a misguidance effect: When the target and non-target features from the previous trial switch, one of the inconspicuous non-targets is frequently selected prior to the target (e.g., Becker, 2008a,b). By contrast, repeating the target and non-target features from the previous trial leads to a high proportion of first eye movements to the target (Becker, 2008a,b; Becker & Horstmann, 2009; McPeek, Maljkovic, & Nakayama, 1999). These results show that priming indeed changes the attentional weights or gains of the target feature (target activation) and/or the non-target feature (non-target inhibition), and does not only modulate the speed of target selection (as would be expected, for instance, on an episodic retrieval view of priming, that repeating the search display merely speeds the retrieval of memory traces without changing attentional gains; Becker, 2008a,b; Hillstrom, 2000).

Originally, the attentional bias was thought to operate on the specific target or non-target feature values, or the corresponding feature maps, and feature priming effects were supposedly driven by intertrial transfers of feature-specific information (e.g., Lamy, Antebi, Aviani, & Carmel, 2008; Leonard & Egeth, 2008; Maljkovic & Nakayama, 1994). However, recent research calls this view into question. According to a feature-based account, priming effects should emerge when either the target changes, or when the non-target features change, compared to the previous trial (e.g., Kristjansson & Driver, 2008; Lamy et al., 2008; Maljkovic & Nakayama, 1994). Moreover, since the target and non-target features are coded independently on separate feature maps, switch costs arising at changes of the target feature and non-target feature should be additive (e.g., Lamy et al., 2008). By contrast, Becker (2010, 2013a) found that both the target and non-target features can change across trials without incurring any switch costs. The critical factor that determines whether switch costs will be present or absent is a change in the feature relationships. If the relative feature of the target remains constant across trials, changing the target and/or non-target features will not incur any switch costs. Switch costs are observed only when the coarse relationship between the target and non-target features reverses across trials; for instance, when the target changed from being larger to being smaller than the non-targets, or vice versa (see Figure 3 for an overview of the results).

Feature relationships have been demonstrated to account for intertrial switch costs in search for colour, luminance, size (e.g., Becker, 2010), and shape (Becker, 2013a). Figure 3 depicts the conditions and results from a colour search task (top panels; Becker, 2010; Exp. 1) and a shape search task (bottom panels; Becker, 2013a; Exp. 1B). In the experiments, observers had to search for a target that differed in colour or shape from the non-targets, and to identify the item inside the target (e.g., "o": right button press; "+": left button press). The colour of the target and/or the non-targets varied differently in 3 blocked conditions, and switch costs were assessed by measuring the proportion of first eye movements to the target ("hits") on repeat trials versus switch trials.

In the *relation same* condition, the colours changed such that the target always had the same feature relationship to the non-targets (e.g., yellower, both on repeat and non-repeat trials). In the *relation reversal* condition, the colours changed such that the relative colour of the target changed on switch trials (e.g., from redder to yellower or vice versa), and in the *full switch* condition, both the relative and the absolute colour of target and non-targets changed, whereby the target adopted the previous non-target colour and vice versa. In the colour search task, the non-target colour always remained the same in the relation same condition and the relation reversal condition, and the conditions differed only in the constancy versus inconsistency of the target's relative colour. As shown in Figure 3A, switch costs were observed only in the relation reversal condition but failed to occur in the relation same condition, in line with a relational account (Becker, 2010).

These and similar results (Becker, 2010) are however still consistent with alternative feature-based explanations. Note that the possible target colours were more similar to each other in the relation same condition (yellow, orange) than in the relation reversal condition (yellow, red). Hence, the findings could also be explained by a grouping account, which assumes that similar colours can be grouped together into a single-target representation (e.g., Duncan & Humphreys, 1989).

Subsequent experiments ruled out these alternative explanations. In one experiment, observers were asked to search for a star shape with a unique number of spikes (Figure 3b, bottom panels). In the relation same condition, the target star had consistently fewer spikes than the non-targets; however, on switch trials, the target adopted the shape of the non-targets from the previous trial or vice versa, so that each switch trial instantiated a half-switch between target and non-target features. In the relation reversal condition, the target was the same on all trials, but the non-target stars changed such that the target was sometimes less spikey, and sometimes spikier, than the non-targets. Despite the fact that both the target and non-target features changed in the relation same condition, no switch costs were found on switch trials. However, in the relation reversal condition, changing only the non-target features across trials led to significant switch costs (Becker, 2013a).

The finding of switch costs in the relation reversal condition but no switch costs in the relation same condition cannot be explained by any of the extant feature-based theories of priming, and provide strong support for a relational account. In line with the relational account, the results indicate that the target feature is usually encoded relative to the features of the non-target context, and attention is subsequently biased towards the relative feature of the target, not its particular feature value. Information about the relative properties of the target (e.g., larger/redder/darker) is then transferred to the next trial and biases attention to items that have the same feature relationships to the other items as the target had to the non-targets on the previous trial (Becker, 2010). This leads to inter-trial switch costs when the target–non-target relationships reverse across trials (e.g., from redder to not redder or vice versa) and facilitates visual selection of the target when it has the same relative feature as on the previous trial—regardless of whether the target or non-target features themselves were repeated or changed.

Importantly, the feature priming effect has usually been attributed to automatic, bottom-up processes (e.g., Maljkovic & Nakayama, 1994). In the following, it is shown that attention is also biased to feature relationships in a top-down controlled manner.

Attention Is Top-down Tuned to Feature Relationships

The top-down selection bias has often been examined in visual search tasks with an additional irrelevant distractor. Observers

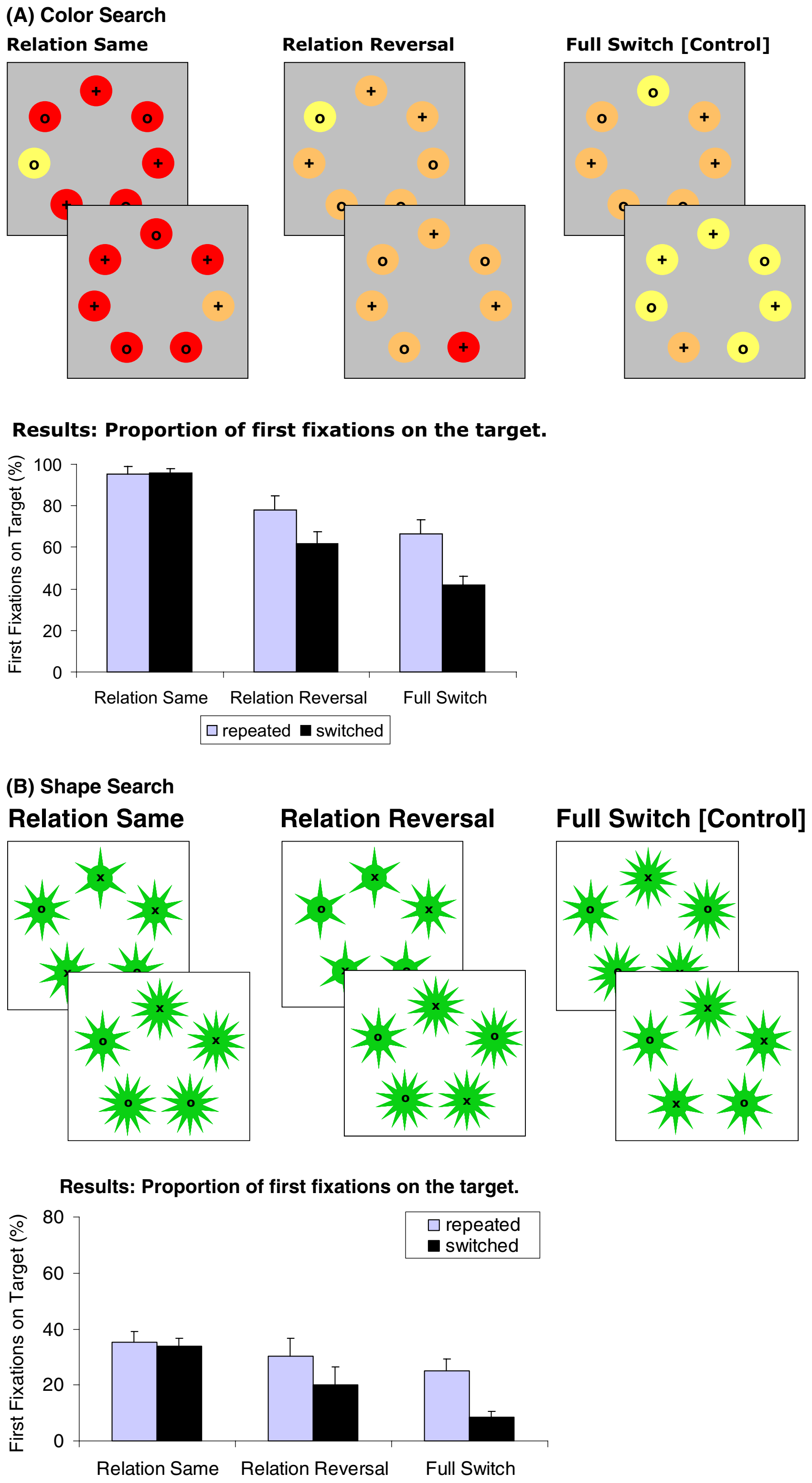

Figure 3 Feature Priming Operates on Feature Relationships. (A) Colour Pop-Out Search. (B) Shape Search (Becker, 2013a; Exp. 1B). *Note*. Examples of the stimulus conditions and the results in the colour search task (A, top panels) and shape search task (B, bottom panels). As shown in the mean proportion of first eye movements to the target, visual selection of the target was not impaired on switch trials in the relation same condition, whereas switch trials produced large and significant costs in the relation reversal condition and the full switch control condition. These results have been replicated with different shapes, including geometrical figures such as triangles and squares (Becker, 2013a) and other features (e.g., size, luminance; see Becker, 2010).

are typically asked to search for a target with a particular feature and to ignore the distractor. To assess how attention was biased to the target feature, the features of the distractor are varied, and it is assessed whether and to what extent it can attract attention ("attentional capture"). Previous studies mostly varied the distractor such that it was either featurally similar or dissimilar from the target (e.g., Ansorge & Heumann, 2003; Becker, Ansorge, & Horstmann, 2009; Folk & Remington, 1998; Folk et al., 1992). The majority of studies found that irrelevant distractors capture attention only when they are similar to the target feature and concluded that attention was biased to the target feature value. For instance, in a study by Ludwig and Gilchrist (2002), observers had to make a fast eye movement to a red target that was presented among three grey non-target items. The target could be presented at one out of four possible locations in the search display. On a portion of trials, an irrelevant red or green distractor was presented at a different (fifth or sixth) location, and observers were instructed to ignore all items presented at these locations. The results showed that observers frequently selected the target-similar, red distractor, whereas the green distractor was selected only very rarely. These results demonstrate that visual selection was biased towards the target colour (red), not towards all coloured items, or the "odd man out" in the search display.[1] The finding that target-similar distractors can capture the attention and the observer's gaze has been obtained with numerous different stimuli and tasks, and is commonly interpreted in support of feature-based theories of attention (e.g., *feature similarity accounts*; Duncan & Humphreys, 1989; Folk & Remington, 1998; Martinez-Trujillo & Treue, 2004; Maunsell & Treue, 2006; Navalpakkam & Itti, 2007; Wolfe, 1994).

However, the findings are also consistent with a relational account of top-down guidance: Note that a target-similar distractor will also have the same feature relationships to the non-targets as the target (e.g., redder). By contrast, a target-dissimilar distractor will usually have markedly different feature relationships (e.g., greener). Hence, all results previously attributed to a feature-specific selection bias could also be due to a selection bias for the relative target attribute (cf. Becker, 2010).

Indeed, studies testing the relational account against a feature-specific view confirmed that capture depends on the feature relationships of the target and distractor rather than their absolute feature values (e.g., Becker et al., 2010, 2013). In one study, the Ludwig and Gilchrist (2002) paradigm was used, but the number of distractor features was increased to test whether attention was indeed biased to the target feature value (Becker, Harris, Venini & Retell, in press). In one block, observers had to make a fast eye movement to an orange target that was embedded among three gold (yellow-orange) non-targets. In another block, the target and non-target features were reversed (i.e., gold target among three orange non-targets). Thus, the target had a constant colour value, and was also redder or yellower than the non-targets within a block. On each trial, an irrelevant distractor was presented at a non-target location that observers were instructed to ignore. The distractor could be either red, orange, gold, or yellow. Capture by the irrelevant distractor was assessed by assessing the proportion of erroneous eye movements to each distractor (see Figure 4 for an example of the display and distractor conditions).

According to a feature similarity account, attention is biased to the target colour value and hence, the target-similar distractor (i.e., orange or gold) should be selected most frequently. According to the relational account, attention should be biased to the relative colour of the target (i.e., *redder* in search for orange; *yellower* in search for gold). As a result, the target-*dissimilar* red distractor should be selected most frequently in search for orange, because it is redder than all other stimuli (including the target), and thus constitute a relatively better stimulus than the target itself. Correspondingly, in search for the gold target, the yellowest (yellow) distractor should capture attention most strongly.

As shown in Figure 4A, the results supported the predictions of the relational account: In search for the orange target, the red distractor was selected most frequently—significantly more frequently than the target-similar orange distractor. In search for the gold target, the yellow distractor was selected more frequently than the target-similar gold distractor (Becker et al., in press, Exp. 1). These results demonstrate that capture by an irrelevant distractor does not depend on its similarity to the target colour, as previously thought (e.g., Folk & Remington, 1998; Ludwig & Gilchrist, 2002; Martinez-Trujillo & Treue, 2004; Wolfe, 1994), but on whether the distractor matches the relative colour of the target.

A subsequent study used a shape search task, to examine whether a shape target would also be encoded relative to the non-target shapes (Becker et al., in press, Exp. 2). In different blocks, observers had to search for a target star with 14 spikes among 10-spike non-target stars (spikier target condition) or a 10-spike target star among 14-spike non-target stars (less spikey target condition). As shown in Figure 4B, the results mimicked the results from the colour search task, with most fixations on the target-dissimilar distractors that had 18 or 6 spikes (Becker et al., in press). According to the relational account, the extreme distractors are so effective in driving attention because they match the relative feature that the target has on the majority of trials (i.e., being the spikiest or least spikey item) while they are spikier than all other items, including the target itself. By contrast, the target-similar distractors are equal to the target, so that attention is equally strongly drawn to the target and the target-similar distractor.

Of note, the experiments replicated the often-reported "similarity effect," as the target-similar distractors captured attention more strongly than the non-target-similar distractors (e.g., Ansorge & Heumann, 2003; Becker et al., 2009; Folk et al., 1992; Folk & Remington, 1998; Ludwig & Gilchrist, 2002). Simultaneously, the results from the more extreme distractors demonstrate that the presumed "similarity effect" does not depend on target-distractor similarity but on whether the distractor matches the relative colour or shape of the target. These results provide strong support for the relational account.

It could be argued that the results are still consistent with a subset of feature-based accounts, as the reddest and yellowest distractors may also have had a larger feature contrast than the target-similar distractors. However, feature contrast is unlikely to account for the observed results: First, note that the red and yellow distractors always had a larger feature contrast than the gold and orange distractors. Yet, the red and yellow distractors

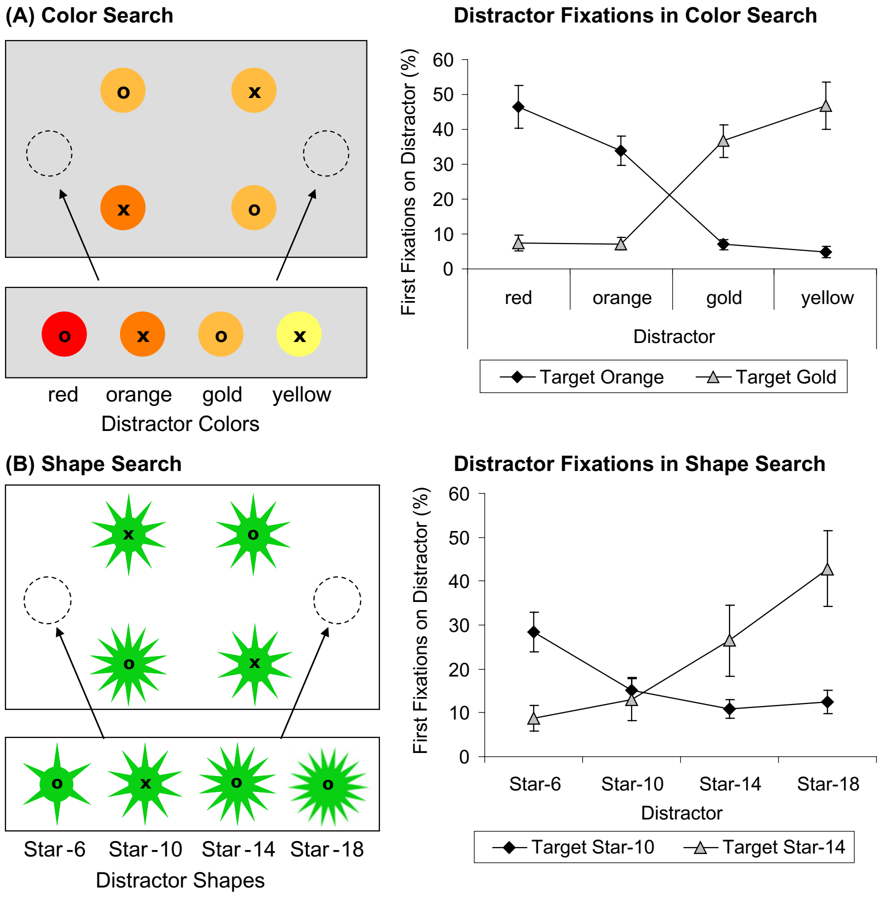

Figure 4 Capture by an Irrelevant Distractor Depends on Feature Relationships. *Note.* In the colour search task (top panels) and the shape search task (bottom panels), observers had to make a fast eye movement to the target and to respond to the item inside the target with a button press. On each trial, one out of the four depicted distractors was presented at one of the two lateral positions to the right or left (dash outlines indicate possible distractor locations, which were never occupied by target). The result graphs (right panels) depict the mean proportion of first eye movements to each of the different distractors, separately for the different target conditions (line graphs).

captured attention only when they matched the relative target colour, not when they mismatched. This result would seem to argue against the view that the distractors attracted more attention because of their higher feature contrasts. Second, feature contrasts presumably do not play a role in shape search tasks; yet, the shape search task showed exactly the same results. Last but not least, the same results were observed in subsequent studies, which used the spatial cueing paradigm to ensure that the feature contrasts of target-similar versus target-dissimilar distractors were equal (e.g., Becker et al., 2010, 2013). In the spatial cueing paradigm, the distractor is presented very briefly prior to the target in a separate cueing display. Hence, the feature contrast of different distractors can be controlled by presenting them among three distractors of a different colour (which constitute the context for the distractor; see Becker et al., 2013). Another advantage of the spatial cueing paradigm is that the colour of a single distractor and its relative colour to the distractor context can be varied independently of the target and non-target colours in the search display. The results of several spatial cueing studies (Becker et al., 2010, 2013) corroborated the findings of the visual search studies described above. An irrelevant distractor captured attention when its relative colour matched the relative colour of the target, regardless of whether it had the same colour as the target or not.

Two studies demonstrated that attention is even biased to feature relationships when the non-targets differ in two opposing directions from the target (Becker et al., 2013, Exp. 2 and 3). In search for an orange target among two red and two yellow non-targets, attention was still biased to the relative colour of the target, to select the item with the intermediate feature value (or the item that was neither the reddest nor the yellowest item in the display). An orange distractor presented prior to the target captured only when it was embedded among redder and yellower items, whereas it failed to capture when it was surrounded by yellower distractors that rendered the distractor the reddest item in the display. Similarly, a target-dissimilar gold distractor captured attention when it was embedded among redder and yellower items, but not when it was the yellowest item in the display. These results suggest that attention can be simultaneously tuned to different feature relationships to select an intermediate colour (cf. Becker et al., 2013). This is an important finding, because it suggests that a more complex relational search strategy can mimic a feature-specific search strategy. Moreover, the results suggest the existence of a mechanism that may be able to extract (quasi-)feature-specific information from information about feature relationships (see also Becker et al., in press; Harris, Remington & Becker, 2013).

Implications for Current Theories of Visual Attention

The results summarised above indicate that attention and eye movements are guided by feature relationships, not by feature-specific information. Moreover, visual selection was biased to the relative feature of the target both when the target feature varied randomly in so-called bottom-up controlled priming effects (e.g., Becker, 2008c, 2010, 2013a) and when the target feature was known and capture by the distractor was presumably contingent on top-down controlled processes (e.g., Becker, 2013a,b; Becker et al., 2010; Harris et al., 2013). These findings question the strict distinction between bottom-up and top-down controlled processes, which is a centrepiece of current models of visual attention (e.g., Folk & Remington, 1998; Lee et al., 1999; Navalpakkam & Itti, 2007; Treisman & Sato, 1990; Wolfe, 1994). However, current theories of visual selection could be quite easily modified to account for both "bottom-up" priming effects and "top-down" attentional biases (e.g., Becker, 2013a; Navalpakkam & Itti, 2006, 2007; Wolfe, Butcher, Lee, & Hyle, 2003).

The finding that attention is preferentially biased to the relative attributes of the target presents a much more serious challenge for the feature detector concept (Becker, 2010, 2013b; Becker et al., 2010, 2013, in press; Harris et al., 2013). The finding that attention is biased to feature relationships across a wide range of stimulus conditions (e.g., colour, luminance, size, and shape) suggests that relational search is a general principle of guidance. This has important theoretical implications: First, the features in the context apparently play a much more important role than previously acknowledged. Even more importantly, the findings imply that visual selection is *not* always independent of the context. Instead, the feature value of the target is encoded relative to the (dominant) features of the context, resulting in a selection bias for objects that match the relative attributes of the target. These results question the feature detector concept, that attention is guided by independently operating feature detectors (e.g., Lee et al., 1999; Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe, 1994).

These findings have important implications for current research, because the feature detector concept has been very influential: The corresponding papers outlining these theories—Treisman and Gelade (1980), Itti, Koch, and Niebur (1998), and Wolfe, Cave, and Franzel (1989) are among the most highly cited papers in vision science and were ranked the 2nd, 38th and 56th most influential papers in vision science, respectively (see also Nakayama & Martini, 2011, for a review). The corresponding visual search theories also greatly influenced applied science, most notably developments in computer vision. As Diamant (2005) remarked, "[w]ith only a minor opposition, the bottom-up/top-down processing principle has been established as an incontestable and dominating leader in both biological and computer vision."

The feature detector concept has exerted its perhaps strongest influence on current neurophysiological research: With few notable exceptions, most studies to date assume that a sensory neuron that responds more strongly to the target than to the non-target *is a* feature detector, and that it plays a role in guiding attention (e.g., Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999). In light of the current evidence, both of these assumptions need to be re-examined. Specifically, future studies need to carefully assess whether and to what extent the response characteristics of the neuron depend on the features of the target versus the features present in the surround (both inside and outside the receptive field; see, e.g., Cavanagh, Joiner, & Wurtz, 2012; Li, Thier, & Wehrhahn, 2001). As mentioned above, the relational account is consistent with many findings that have previously been interpreted in support of feature-specific selection accounts (e.g., Becker, 2010). To understand visual search, we have to understand how attention is biased to feature relationships on the neuronal level.

Conclusion

The concept that visual attention is guided by feature detectors, originally derived from the Nobel prize winning work of Hubel and Wiesel (1969), has become the centrepiece of current theories of visual attention. Yet, new findings question the assumption that attention is guided by the output of neurons that respond to elementary features. Several studies demonstrate that attention is guided in a relational or context-dependent manner, suggesting the need to modify the architecture of current theories of visual attention.

Note

1. Ludwig and Gilchrist (2002) also tested an onset distractor to assess whether top-down controlled processes can dominate selection over bottom-up saliency signals. These results were not discussed, as they are outside the scope of the present article.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Science*, *8*, 457–464. doi:10.1016/j.tics.2004.08.011

Ansorge, U., & Heumann, M. (2003). Top-down contingencies in peripheral cueing: The roles of color and location. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 937–948. doi:10.1037/0096-1523.29.5.937

Baldassi, S., & Burr, D. C. (2000). Feature-based integration of orientation signals in visual search. *Vision Research*, *40*, 1293–1300. doi:10.1016/s0042-69890000029-8

Becker, S. I. (2008a). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research*, *48*, 664–684. doi:10.1016/j.visres.2007.10.025

Becker, S. I. (2008b). The mechanism of priming: Episodic retrieval or priming of pop-out? *Acta Psychologica*, *127*, 324–339. doi:10.1016/j.actpsy.2007.07.005

Becker, S. I. (2008c). Can intertrial effects of features and dimensions be explained by a single theory? *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1417–1440. doi:10.1037/a0011386

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. doi:10.1037/a0018808

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

Becker, S. I. (2013b). Guidance of attention by feature relationships: The end of the road for feature map theories? In M. Horsley, M. Eliot, R. Knight, & B. Riley (Eds.), *Current trends in eye tracking research*. Springer.

Becker, S. I., Ansorge, U., & Horstmann, G. (2009). Can intertrial priming account for the similarity effect in visual search? *Vision Research*, *49*, 1738–1756. doi:10.1016/j.visres.2009.04.001

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:10.1037/a0020370

Becker, S. I., Folk, C. L., & Remington, R. W. (2013). Involuntary attentional capture does not depend on feature similarity, but on target-nontarget relations. *Psychological Science*. doi:10.1177/0956797612458528

Becker, S. I., Harris, A. M., Venini, D., & Retell, J. D. (in press). Visual search for colour and shape: When is the gaze guided by feature relationships, when by feature values? *Journal of Experimental Psychology: Human Perception and Performance*.

Becker, S. I., & Horstmann, G. (2009). A feature weighting account of priming in conjunction search. *Attention, Perception, & Psychophysics*, *71*, 258–272. doi:10.3758/APP.71.2.258

Cavanagh, P., Joiner, W. M., & Wurtz, R. H. (2012). Suppressive surrounds of receptive fields in monkey frontal eye field. *The Journal of Cognitive Neuroscience*, *32*, 12284–12293. doi:10.1523/JNEUROSCI.0864-12.2012

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393–404. doi:10.1016/S0042-69890200596-5

Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*, 1–13. doi:10.3758/BF03195009

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

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology*, *62*, 1716–1722. doi:10.1080/17470210902811249

Diamant, E. (2005). Does a plane imitate a bird? Does computer vision have to follow biological paradigms? In M. DeGregorio, V. Di Maio, M. Frucci, & C. Musio (Eds.), *BV&AI 2005, LNCS 3704* (pp. 108–115). Berlin Heidelberg: Springer-Verlag.

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

Folk, C. L., & Remington, R. W. (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. doi: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. doi:10.1037/h0036324

Foster, D. H. (2003). Does color constancy exist? *Trends in Cognitive Science*, *7*, 439–443. doi:10.1016/j.tics.2003.08.002

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, *9*, 1–13. doi:10.1167/9.11.1

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

Hillstrom, A. P. (2000). Repetition effects in visual search. *Perception & Psychophysics*, *62*, 800–817. doi:10.3758/BF03206924

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human awareness. *Science*, *317*, 823–825. doi:10.1126/science.1143515

Hubel, D. H., & Wiesel, T. N. (1969). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*, 215–243.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention and rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*, 1254–1259. doi:10.1109/34.730558

Jacoby, O., Kamke, M. R., & Mattingley, J. M. (2013). Is the whole more than its parts? Estimates of average size and orientation are susceptible to object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 233–244. doi:10.1037/a0028762

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*, 751–761. doi:10.1016/S0896-62730080734-5

Kristjansson, A., & Driver, J. (2008). Priming in visual search: Separating the effects of target repetition, distractor repetition and role reversal. *Vision Research*, *48*, 1217–1232. doi:10.1016/j.visres.2008.02.007

Lamy, D., Antebi, C., Aviani, N., & Carmel, T. (2008). Priming of pop-out provides reliable measures of target activation and distractor inhibition in selective attention. *Vision Research*, *48*, 30–41. doi:10.1016/j.visres.2007.10.009

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

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. doi:10.1080/13506280701580698

Li, W., Thier, P., & Wehrhahn, C. (2001). Neuronal responses from beyond the classic receptive field in V1 of alert monkeys. *Experimental Brain Research*, *139*, 359–371. doi:10.1007/s002210100757

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. doi: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:10.3758/BF03209251

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. doi:10.1016/j.cub.2004.04.028

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

McPeek, R. M., Maljkovic, V., & Nakayama, K. (1999). Saccades require focal attention and are facilitated by a short-term memory system. *Vision Research*, *39*, 1555–1566. doi:10.1016/S0042-69899800228-4

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*, 1526–1537. doi:10.1016/j.visres.2010.09.003

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

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

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*, 739–744. doi:10.1038/89532

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536. doi:10.1038/323533a0

Sweeny, T. D., Grabowecky, M., Paller, K., & Suzuki, S. (2009). Within-hemifield perceptual averaging of facial expression s predicted by neural averaging. *Journal of Vision*, *9*, 1–11. doi:10.1167/9.3.2

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606. doi:10.3758/BF03211656

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

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. doi:10.1016/0010-02858090005-5

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

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 483–502. doi:10.1037/0096-1523.29.2.483

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433. doi:10.1037/0096-1523.15.3.419