Cognition 260 (2025) 106132

Contents lists available at ScienceDirect

Cognition

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

Visual search is relational without prior context learning

Stefanie I. Becker [*], Zachary Hamblin-Frohman, Koralalage Don Raveen Amarasekera

School of Psychology, The University of Queensland, Brisbane, Australia

ARTICLE INFO

Keywords:
Attention
Visual search
Eye movements
Distractor
Relational account
Relative features

ABSTRACT

The most prominent models of visual attention assume that we tune attention to the specific feature value of a sought-after object (e.g., a specific colour or orientation) to aid search. However, subsequent research has shown that attention is often tuned to the *relative* feature of the target, that the target has in relation to other items in the surround (e.g., redder/greener, darker/lighter, larger/smaller), in line with a Relational Account of Attention. Previous research is still limited though, as it used repeated-target designs and relatively sparse displays. With this, it is still unknown whether we can indeed tune attention to relative features prior to the first eye movement, or whether this requires context knowledge gained from experience. Moreover, it is unclear how search progresses from one item to the next. The present study tested these questions in a 36-item search display with multiple distractors and variable target and non-target colours. The first fixations on a trial showed that these displays still reliably evoked relational search, even when observers had no knowledge of the context. Moreover, the first five fixations within a trial showed that we tend to select the most extreme items first, followed by the next-extreme, until the target is found, in line with the relational account. These findings show that information about the relative target feature can be rapidly extracted and is used to guide attention in the first fixation(s) of search, whereby attention only hones in on the target colour after multiple fixations on relatively more extreme distractors.

1. Introduction

It is well-known that we cannot consciously process all objects in a visual scene at once. Visual attention selects objects for in-depth processing, often by guiding our gaze to relevant parts in a scene (e.g., Deubel & Schneider, 1996). Much effort has been devoted to determine which items in a scene will be attended first, and more generally, to identify the processes involved in creating our rich mental representation of the visual environment (for a review, see Carrasco, 2011; Wolfe, 2021).

To date, it is widely accepted that attention can be guided by both, bottom-up, stimulus-driven processes and top-down, goal-driven processes (e.g., Wolfe, 2020). For example, attention can be reflexively drawn to visually salient events such as a bright flash, a movement, or the sudden appearance of an object (e.g., Theeuwes, 2004, 2013), or it can be top-down tuned to select items with certain attributes (e.g., colours: red, green) to help goal-related behaviours such as finding a friend in a crowd (e.g., Desimone & Duncan, 1995; Folk & Remington, 1998; Wolfe, 1994, 2021). Correspondingly, current models of visual attention typically include both a bottom-up and a top-down component to predict which items in a visual scene will be selected first (e.g., Wolfe, 1994, 2021). Top-down tuning is typically modelled as an increase in the firing rate of sensory neurons in response to specific stimulus attributes (e.g., red, green; Navalpakkam & Itti, 2007; Yu, Hanks, & Geng, 2022). For example, when looking for an orange in a fruit basket, we would tune attention to the specific colour of orange, which increases the output of neurons that respond to orange and prioritises colour-matching items for selection.

It is commonly assumed that attention is tuned to the feature value of an object that a person is looking for (e.g., particular shade of orange; e.g., Duncan & Humphreys, 1989; Navalpakkam & Itti, 2007). However, to date, there are also several accounts of non-veridical tuning. Navalpakkam and Itti (2007) noted that tuning attention to the exact target feature value would not be beneficial when the target is very similar to surrounding irrelevant non-target items, as tuning attention to, for example, orange, would also boost the response gain of similar colours (e.g., red-orange or yellow-orange), leading to a poor signal-to-noise ratio (SNR). They proposed that attention would be tuned to a feature value that is slightly shifted away from similar nontargets, to increase the signal-to-noise ratio (SNR; e.g., tuning to yellow-orange, when an orange target is presented among red-orange items; Navalpakkam & Itti, 2007). According to their *optimal tuning account*, attention is always

---

[*] Corresponding author.

https://doi.org/10.1016/j.cognition.2025.106132
Received 13 May 2024; Received in revised form 25 March 2025; Accepted 27 March 2025

tuned to the feature value that maximises the ability to discriminate the target from the non-targets (i.e., maximise the SNR). Thus, attention would only be tuned to the exact target feature value when the target is presented alone or among dissimilar other items. When the target object is surrounded by other similar objects and discrimination is difficult, attention should be shifted to a slightly exaggerated target feature value, away from the non-target feature.

In line with this idea, Navalpakkam and Itti (2007) showed that participants erroneously selected a slightly shifted non-target colour instead of the target on perceptual probe trials, when the target in the intermixed visual search trials was consistently embedded among similarly coloured non-targets in a prior visual search. For example, when participants searched for a greenish-blue target among slightly more bluish non-targets, they selected a slightly more yellowish-green colour as the target colour in intermixed perceptual probe trials (Navalpakkam & Itti, 2007; see also Geng, di Quattro, & Helm, 2017; Scolari, Byers, & Serences, 2012).

Another account of non-veridical tuning is the Relational Account, which proposes that attention is not tuned to a specific feature value but a *relative* property of the target object (e.g., redder, darker, larger; Becker, 2010). In natural visual environments, tuning attention to a particular feature value may not be beneficial, as specific feature values such as the size, shape and colour of objects vary greatly with differences in distance, perspective and shading. To allow efficient selection of the targets in these noisy environments, attention could be tuned in a context-dependent manner to objects, biasing attention to a feature that the target has *relative* to the other items in the context (e.g., reddest, largest, darkest; Becker, 2010). For example, when searching for an orange in a fruit basket, the visual system would quickly assess the distribution of colours in the visual scene to determine the dominant colours, and tune attention to the relative colour that would best discriminate the target from the dominant coloured non-target items. Thus, when the fruit basket (or visual scene) contains many yellow or green objects, attention would be tuned to all redder items or the reddest item, whereas attention would be tuned to all yellower items or the yellowest item if the fruit basket (or visual scene) contains mainly red objects. As a consequence of this broad relational tuning, the item that maximally fulfils the relevant feature relationship will be selected first (e.g., reddest item). Hence, when attention is tuned to all redder items, the reddest item in the visual field should be selected first, followed by the next-reddest, and so forth.

In line with this prediction, several visual search studies showed that when an orange target is presented among mostly yellow(er) items, a red irrelevant distractor was more likely to be attended first, even when it was quite dissimilar from the target, suggesting that attention was biased to all redder items, or the reddest item (e.g., Becker, 2010; Becker, Folk, & Remington, 2013; Hamblin-Frohman & Becker, 2021; York & Becker, 2020). Selection of the distractor was reflected both in a high proportion of first eye movements to the red distractor and elongated response times (RTs) in the presence of a red distractor. Several studies also included a visually salient distractor with a dissimilar colour (e.g., blue), but found no or only very weak effects of saliency, ruling out that selection of the relatively matching (e.g., red) distractor was mediated by bottom-up, stimulus-driven processes (e.g., Martin & Becker, 2018; York & Becker, 2020).

Both optimal tuning and the relational account propose that attention can be tuned to features in a non-veridical manner, but the theories propose very different underlying mechanisms. Fig. 1 depicts the different tuning functions for the ‘standard’, feature-specific account, optimal tuning and relational tuning when searching for an orange target.

As shown in the Figure, the main difference between optimal tuning and the relational account is that optimal tuning predicts that attention is tuned to an ‘exaggerated’ feature value that is shifted away from the nontargets (e.g., Navalpakkam & Itti, 2007), whereas the relational account predicts tuning to all redder items or the reddest item (e.g., Becker, 2010). This can result in attentional capture by colours that are very dissimilar from the target colour (e.g., full red in search for orange; see Fig. 1; Becker, 2010).

Two previous studies combined the two paradigms used by optimal and relational tuning accounts to determine whether attention is tuned to all relatively matching features or an optimally shifted feature value (Hamblin-Frohman & Becker, 2021; Yu et al., 2022). The results showed that attention is biased to all relatively matching items, indicating that early visual selection (i.e., which item is selected first) follows the predictions of the relational account. However, target identification judgements were skewed towards the slightly shifted feature value, showing that perceptual decision-making (i.e., decisions about whether a selected item is the target or not) follows the optimal tuning account (Hamblin-Frohman & Becker, 2021; Yu et al., 2022). Thus, visual attention can be tuned rather broadly towards the target’s relative feature, leading to frequent selection of very dissimilar colours that only share the target’s relative colour, whereas perceptual decision-making is tuned rather sharply towards an exaggerated target feature value, so that only very similar colours are mistaken for the target colour. The finding that attention is tuned towards the relative properties of the target, whereas optimal tuning describes later processes of perceptual decision-making indicates important progress in attention research (Hamblin-Frohman & Becker, 2021; Yu et al., 2022). However, there are still significant empirical gaps.

Of note, the relational account makes two important claims that have never been tested; viz:

1. In relational search, we first select the most extreme (relatively matching) item first (e.g., reddest item in search for redder), then the next-extreme (e.g., next-reddest) item, etc., until we find the target feature.
2. Relational search does not require prior knowledge of the context: The visual system can quickly assess the dominant feature in a visual scene, and tune attention to the relative feature of the target prior to selecting any items.
   a. Relational search is a *default* search strategy that will be used whenever the target differs from a clear majority of items in a single direction.

2. Selection mechanism and measures

It is widely acknowledged that, in a visual search task, the search items can compete for visual attention. This competition can be modulated by bottom-up factors (e.g., saliency), top-down tuning and the trial history (which comprises both learning and short-term priming effects; e.g., Awh, Belopolsky, & Theeuwes, 2012; Becker, Grubert, Ansorge, & Horstmann, 2023). All of these factors can guide visual attention, whereby the relative importance of each factor depends on the specifics of the task and the environment (including the variability vs. constancy of the stimuli; e.g., whether the target or an irrelevant salient distractor are repeated or not; e.g., Becker, Grubert, et al., 2023; see also Becker, 2018). Previous studies assessing the attentional tuning function have often used eye movements to assess how attention was top-down tuned to the target (e.g., Hamblin-Frohman, Chang, Egeth, & Becker, 2022; Navalpakkam & Itti, 2006; Yu et al., 2022, 2023). Eye movements and attention are tightly coupled, in that an eye movement to a location is always preceded by an attention shift to the same location (Deubel & Schneider, 1996; Peterson, Kramer, & Irwin, 2004). Thus, monitoring eye movements allows drawing inferences about where attention was

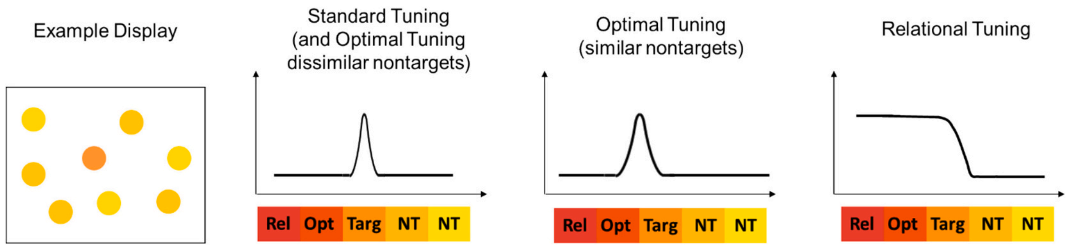


Fig. 1. Overview of different theoretical tuning functions reflecting how attention would be tuned to the target in the example display (left), which shows an orange target (Targ) among yellow-orange and yellow non-targets (NT). The standard view (2nd from left) is that attention would be tuned to the target colour (Targ). According to Optimal Tuning (3rd from left), attention would be tuned to a slightly exaggerated, 'optimal' target feature value that is shifted away from the nontargets (here: a more reddish orange), to increase the SNR (optimal colour; Opt). According to the Relational Account (right), attention would be tuned to the relative colour of the target, that the target has relative to the other items in the surround (here: redder), which can include vastly dissimilar colours (relatively matching colour, Rel). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

allocated at any point in time.[1] Previous studies have mainly used the first eye movement in a trial to measure how attention was biased to the items in the field, as the first eye movement is not contaminated by later, post-selective processes such as distractor rejection, decision-making or response selection, that can all come into play after the first selection (e.g., Becker, 2010; Becker, Retell, & Wolfe, 2023; Horstmann & Becker, 2020; Zelinksy & Sheinberg, 1997).

According to the relational account, when attention is tuned to the relative target feature (e.g., redder), all items that differ from the target in this direction should have a higher attention-driving capacity than the target itself (e.g., Becker, 2010). If we draw a line through feature space (e.g., CIE colour space) connecting the target and dominant non-target feature, all items that are located on this line further away from the target should attract attention, and theoretically, more strongly than the target itself (if presented together with the target). The feature distance or similarity/dissimilarity from the target should not matter.

However, in practice, an item that is further located away from the target is more readily identifiable as having the correct relative properties (e.g., is in fact redder) than an item that is more similar to the target and potentially confusable with the target, which can explain why we have occasionally observed stronger capture by more dissimilar distractors that are further away from the target than a more similar distractor (e.g., Becker, Valuch, & Ansorge, 2014a, 2014b; York & Becker, 2020). At the same time, items that are located further away from the target are more likely to deviate from the line in feature space that defines the boundaries for target-matching relative features and/or can be mixed with other, visibly different features (e.g., violet), which would reduce their attention-driving capacity (e.g., York & Becker, 2020). Relational search transcends standard categorical boundaries in colour space (e.g., Yu et al., 2022), but selection will be reduced when a distractor ceases to be more extreme than the target with respect to the relative feature (e.g., when a violet distractor is not redder than the orange target).

The principal selection mechanism of the relational account is that we should first select the most extreme, relatively matching item first, followed by the next-extreme relatively matching item and so forth, until the target feature is selected (e.g., Becker, 2010). For example, in search for an orange, redder target we should first select the reddest item (e.g., full red), followed by the next-reddest item (e.g., red-orange), etc., until we find the target colour (e.g., Becker, 2010). This assumption is central to the relational account and sets it apart from other feature-based views such as a linear separability view (e.g., Bauer, Jolicoeur, & Cowan, 1996; D'Zmura, 1991) or a feature divider account (e.g., Huang & Pashler, 2005).[2] However, this assumption has never been tested, as previous studies only contained a single differently coloured distractor (e.g., Becker, 2010; Becker et al., 2013; Becker, Harris, York, & Choi, 2017; Becker, Valuch, & Ansorge, 2014a, 2014b; Hamblin-Frohman & Becker, 2021).

To test this hypothesis, it would be necessary to assess how visual search progresses through multiple differently coloured distractors and track the participants' eye movements over multiple successive fixations, which is one of the major aims of the present study.

3. Pre-requisites for top-down (relational) tuning

A second important hypothesis of the relational account is that biasing attention to the relative feature of the target does not require pre-knowledge of the context or on-task learning. Rather, the visual system can quickly assess the distribution of features present in a visual scene and use this information to bias attention in a context-dependent fashion to the target. For example, if we are looking for an orange target among mostly yellow other items, the visual system can quickly extract the dominant feature in the visual scene ("yellow") and use this information to tune attention to the relative feature of the target that allows best discriminating it from the non-targets (i.e., redder).

Previous studies on feature averaging and scene gist perception already showed that the visual system can quickly assess the distribution of features present in a visual scene and report the average feature in a purely automatic fashion (e.g, Chong & Treisman, 2005a, 2005b; Oriet & Brand, 2013), as well as quickly extract the gist of a scene and accurately categorise it (e.g., as a dinner vs. picnic scene; Oliva & Torralba, 2006; Wyble, Folk, & Potter, 2013). As Rosenholtz (in press) remarked, this remarkable ability to quickly extract the summary statistics of a scene is poorly integrated in most current models of visual attention. However, according to the relational account, this mechanism is important for extracting information that is essential for successfully tuning attention to the relative feature of the target (e.g., Becker, 2013a, 2013b).

Biasing attention to the relative feature of the target rather than its absolute feature (e.g., the exact colour, brightness or size) is thought to convey an advantage in natural environments, where the absolute feature values of an item vary a lot with differences in perspective, distance and shading (e.g., Becker, 2013a, 2013b; Grössle, Schubö, & Tünnermann, 2023).

Importantly, evaluation of the dominant features in a visual scene

---

[1] It is possible to move attention while the eyes remain fixated. However, shifting attention without a concomitant eye movement will still take time and lead to delays in executing eye movements. This would be reflected in longer saccadic latencies, so that covert attention shifts remain detectable in eye movement paradigms.

[2] Previous studies have shown that the relational account can potentially explain the linear separability effect (e.g., Becker, 2010; Brand, Oriet, Johnson, & Wolfe, 2014). However, the reverse is not true, as a linear separability account cannot explain stronger capture for relatively better-matching distractors than the target, or the dynamics of feature priming effects (e.g., Becker, 2010; Becker, Valuch & Ansorge, 2014).

and biasing attention to the relative feature of the target is supposed to occur automatically and prior to selecting any stimuli. Thus, it should be possible for observers to tune attention to the relative feature of the target even without prior knowledge of the context. Visual search may improve with learning and on-task experience, but this should not be critical for successfully selecting an item on the basis of its relative features.

Even though this is an important tenet of the relational account that also sets it apart from other theories (e.g., optimal tuning; Navalpakkam & Itti, 2007), it has never been tested. In previous studies, the target and non-targets were typically repeated numerous times (e.g., Becker, 2010). This leaves open the possibility that relational search was based on learning the target's discriminative relative feature, or that the visual system adapted weight settings over consecutive trials in a trial-and-error fashion to bias attention to the relative feature. Hence, it is still an open question whether the visual system can indeed quickly evaluate the context to bias attention to relative features prior to the first eye movement.

Another, related limitation of previous research is that it was never tested whether the visual system is indeed sensitive to the dominant feature in the context. Previous research has shown that attention will be biased to the relative feature of the target (e.g., redder) when the target can be selected in virtue of its relative feature as the first item on a majority of trials (such as in variants of all pop-out search tasks; e.g., when searching for an orange target among yellow-orange and yellow other items; Becker, 2010). By contrast, attention will be tuned to the specific feature value of the target when the target does not have a relatively extreme feature value and is sandwiched between other colours (e.g., orange target among red and yellow other items; Becker, Harris, Venini, & Retell, 2014). However, in studies demonstrating relational search, the target always had a relatively extreme feature value on more than half of all trials (e.g., reddest), leaving it an open question whether we would also observe relational search when the context contains a dominant feature but the target never has an extreme feature value (e.g., when searching for an orange target among many yellow non-targets, when there are always a few red distractors present). It is conceivable that attention would be biased to the specific feature value of the target in these conditions (e.g., orange), which would invalidate the claim that the visual system quickly assesses the dominant feature in the context and biases attention to the relative feature of the target.

4. Overview of study

The aim of the present study was to test the two main untested assumptions of the relational account in search displays that always contained thirty-six items and included multiple different distractors that could match or mismatch the relative feature of the target (see Fig. 2). For example, in the redder target condition, an orange target was presented among 29 yellow-orange and yellow nontargets, one distractor with a relatively matching colour (red), two distractors with an optimal colour; i.e., an exaggerated target colour that was slightly shifted away from the nontarget colours (red-orange), two distractors with the target colour (orange), and a saliently different distractor with an unrelated colour (e.g., a blue; see Fig. 2). These displays allows assessing how attention is tuned to the target when the target differs from the majority of non-targets in a single direction (redder) but is never the relationally maximal item (e.g., reddest, bluest, greenest or yellowest item). The displays also allowed assessing the progression of visual search, viz., the hypothesis that the reddest item would be selected first, followed by the next-reddest item, and the target (e.g., Becker, 2010).

The displays also always included a saliently different distractor in each display that was very dissimilar from all other colours, to assess possible bottom-up saliency effects. A similar salient distractor control condition was also included in previous studies and usually showed no or only very weak attention-driving capacity (e.g., Hamblin-Frohman & Becker, 2021; York & Becker, 2020; Yu et al., 2022).

However, Wang and Theeuwes (2020) pointed out that the salient distractor may not have been very salient in the sparse displays used in these studies, which typically contained four to eight items. Hence, it is possible that previous work underestimated bottom-up saliency effects in selection (e.g., Becker, 2010; Becker, Harris, et al., 2014; Hamblin-Frohman & Becker, 2021; Navalpakkam & Itti, 2007; Yu et al., 2022). The present displays allow re-examining possible saliency effects in more ecologically valid conditions.

We also included a manipulation to test whether attention could be tuned to the relative properties of the target without prior training or knowledge of the context. To that aim, we created four blocked conditions in which the search colours varied (bluer, greener, yellower or redder target; see Fig. 2). Participants completed ten trials in each block, and then switched to a different search colour. Prior to each block, participants were only informed about the exact target colour (teal or orange), but not about the colour of the non-target items or the target's relative colour (redder, yellower, greener, or bluer). Thus, analysing the

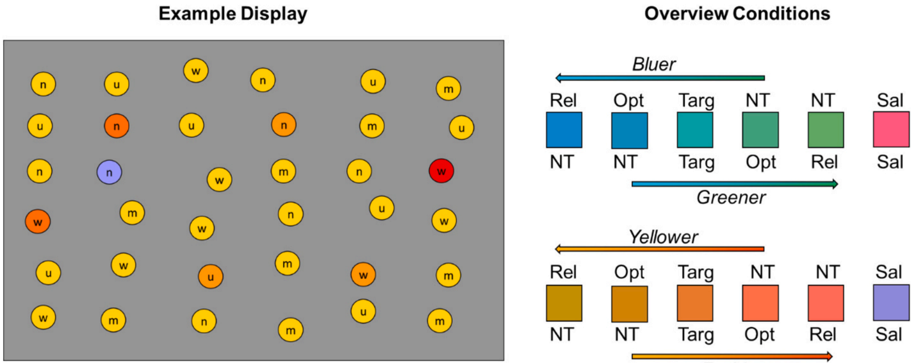


Fig. 2. Left: Example of a visual search display. Search displays consisted of three target-coloured items (here: orange) presented among 29 nontarget coloured items that could have two different colours (here: yellow-orange or yellow). Participants were instructed to search for a target-coloured item (here: orange) that contained the letter *u* or *m*, and to report whether the target contained the u or m (whereas the other orange items contained the letters *w* or *m*). To assess how attention was guided in search, we assessed eye movements to items with a relatively matching colour (here: red), optimal colour (here: red-orange), target colour, and a saliently different colour (here: violet). Right: Colours used to create the four blocked conditions. The target could be teal or orange, and the teal target could be presented among other green-ish non-targets (bluer target), blue-ish non-targets (greener target), while the orange target could be among red-ish non-targets (yellower target) or yellow-ish non-targets (redder target). Prior to a mini-block of 10 trials, participants were informed of the target colour, but not of the colour of the non-targets, which determined whether the target would be relatively redder/yellower or greener/bluer than the non-targets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

first trial of each block allows assessing how attention was tuned to the target under conditions of uncertainty and in the absence of training effects.

To assess how attention was tuned to the target, we measured eye movements to each of the different item types (target and distractors). Specifically, to tap into early processes of visual selection, we analysed the first eye movements on a trial, which are not influenced by prior fixations or search processes.

If attention is tuned to the exact target colour, as proposed by standard feature-specific tuning accounts, we would expect a large proportion of first eye movements on the target-coloured items (orange, teal), regardless of the context colour, and only a few eye movements on the differently coloured distractors (optimal, relational; see Fig. 3a), with declining selection rates for distractors that are less similar to the target. On the other hand, if attention is tuned to the optimal colour, we would expect most first eye movements to the optimal colour and significantly fewer eye movements to other-coloured distractors (see Fig. 3b). Moreover, this effect should only develop over time, as tuning to the optimal feature requires learning (e.g., Navalpakkam & Itti, 2007). According to the relational account, we would expect most first eye movements to the distractor that best matches the relative colour of the target, viz., the distractor that is the reddest, yellowest, greenest, bluest in the visual field, followed by the optimal distractor (as this is the next-maximal, e.g., next-reddest item), and the target-coloured distractors (see Fig. 3c).

Moreover, if relational search is based on a fast, automatic assessment of the dominant colour in the display, we would expect to see these results immediately in the first trial of each block, prior to learning and knowledge of the relative values.

On the basis of previous studies, we expected the salient distractor to attract attention only very weakly (e.g., Becker, 2018; Martin & Becker, 2018). Moreover, as previous studies have shown that participants can quickly learn to ignore or inhibit the distractor, we would expect saliency effects to quickly decrease with training (i.e., over the course of a block; e.g., Gaspelin & Luck, 2018; Hamblin-Frohman et al., 2022).

5. Method

5.1. Participants

To estimate the required sample size, we examined the ability to detect relational vs. feature-specific search in the first fixations on distractors in previous work (Becker, Harris, et al., 2014). The weakest effect was the feature-specific effect ($t(14) = 2.5$, $p = .024$; Becker, Valuch, and Ansorge, 2014b; Exp. 3). The BUCSS tool suggested a target sample size of $N = 32$ to achieve a power of 85 % (with 50 % assurance; Anderson, Kelley, & Maxwell, 2017).

Thirty-four paid participants from The University of Queensland participated in the experiment. Two participants were excluded for having a low search accuracy (< 70 %), leaving 32 participants in the final analysis (M age = 23.1 years ($SD = 1.9$), 24 female). The study was approved by the University's ethics board, and all procedures were in line with the Declaration of Helsinki.

5.2. Apparatus

Stimuli were presented on a 21-in. CRT monitor with a refresh rate of 85 Hz. A chin and headrest were used to hold the participant's heads 600 mm from the screen. Gaze location was measured by an SR-Research Eyelink-1000 eye tracker at 500 Hz sampling rate. The experiment was controlled by Python's PsychoPy (Peirce, 2007).

5.3. Stimuli

All stimuli were presented against a grey background. Each search array contained thirty-six coloured circles (radius: 0.48°) arranged in a six-by-six grid format (see Fig. 2). Stimulus locations were initially selected to have a 4.96° horizontal separation and 4.30° vertical separation (centre-to-centre), which varied because the location of all non-target stimuli was randomly jittered by ±1.43° horizontally and vertically on each trial. The relevant distractors (relational, optimal, target-similar and salient) and the target were not jittered to ensure precise eye tracking and to ensure that these stimuli were never too close to each other. The location of all items was randomly chosen on each trial, with the restriction that the items of interest could never appear in the corner and corner-adjacent positions of the search array (as these positions were too far from fixation), and never in the central four (3.82° from fixation) positions (as these were too close to fixation).

Each coloured circle contained a letter, either 'u', 'n', 'm' or 'w' (height: 0.29°). The target stimulus always contained a 'u' or 'm'. The distractor items of interest never contained the target response characters (u or m). Colours were selected from an equiluminant (30 ± 2 cd/m^2) RGB colour set (see Fig. 2). There were two potential target colours, orange (RGB: [227, 124,52]) and teal (RGB: [56,171,146]), that

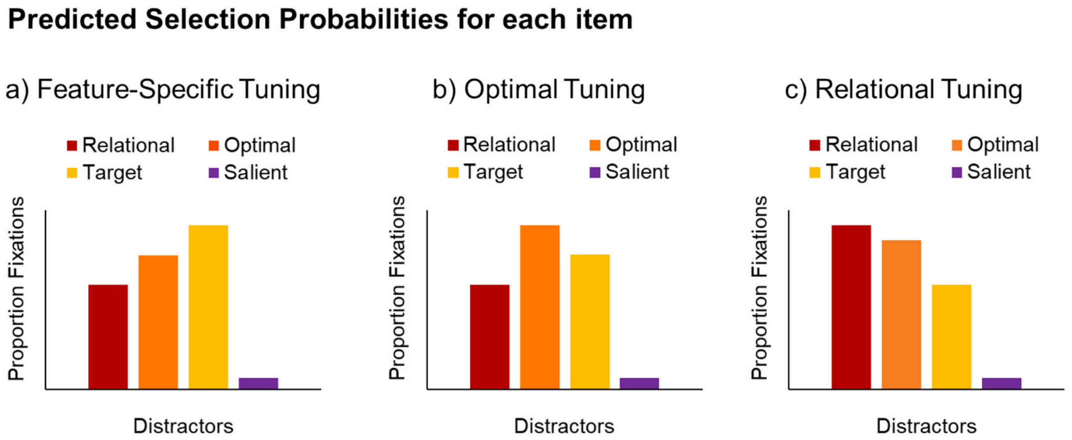


Fig. 3. The predicted selection probabilities for a relational, optimal, target coloured and saliently different distractor (as reflected in the proportion of first fixations on each item), when attention is tuned to a) the specific target colour, b) the colour that optimally distinguishes between the target and the similar non-targets (adapted from Navalpakkam & Itti, 2007), or c) the relative colour of the target. Note that optimal tuning only predicts tuning to a slightly shifted, exaggerated colour when the non-targets are very similar to the target. Both the standard feature-specific view and optimal tuning predict that selection rates should be higher for target-similar and optimal colours than for a relatively matching colour that is quite dissimilar from the target (e.g., red in search for orange). Only the relational account predicts higher selection rates for relatively matching, dissimilar colours than for target-matching colours. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

alternated between blocks. The orange target could either appear among a set of redder non-targets, creating a *yellower* target condition, or among yellower non-targets, creating a *redder* target condition. The teal target could be presented either among greener or bluer non-targets, creating *bluer* or *greener* target conditions, respectively. Each trial contained the same amount of distractor and non-target items: One distractor with a relationally matching colour (R), two distractors with an optimal colour (O), two target-matching distractors (T), and a salient distractor (S) that always had a colour from the opposite side of colour space (e.g., a pink distractor for when the target was teal). The stimuli were always presented among 29 non-target items that were selected from two other colours (e.g., two blue-ish colours; see Fig. 2).

5.4. Design

The colours of the target, distractors and non-targets were always repeated within a mini-block of 10 trials, and mini-blocks alternated between orange and teal targets, with the direction of search (redder/yellower or greener/bluer condition) determined randomly at the start of each block. Participants completed 64 mini-blocks, for a total of 640 trials. The first four blocks were treated as practice trials, leaving 600 trials for the final analysis.

5.5. Procedure

Prior to the experiment, participants were instructed to locate the target-coloured stimuli containing the character *u* or *m* and respond with the corresponding keyboard key as quickly and accurately as possible. Moreover, prior to each mini-block, participants were shown the colour of the upcoming visual search target (either orange or teal). Importantly, no information about the colour of the non-targets was provided; hence participants did not know the relative colour of the target prior to the first trial (i.e., whether it was redder/yellower or greener/bluer).

To ensure stable and accurate eye tracking, participants were calibrated with a randomised 9-point calibration at the beginning of the experiment and we implemented a fixation control prior to each trial. If participants maintained fixation for 650 ms on the central fixation cross (within 2.0° from the centre) prior to the trial, the search array was presented until a manual keypress response (u, m) was recorded. Otherwise, participants were re-calibrated and the trial was started anew (with the fixation control). If an incorrect response was recorded an “Incorrect Response” feedback was displayed for 750 ms. After each trial, a blank grey screen was presented for 750 ms, and the next trial commenced again with the fixation control. For the purpose of the analyses, fixations were always assigned to the nearest stimulus.

6. Results

Overall, accuracy on the letter identification task was high (>90 %). Trials with incorrect responses were excluded from all analyses (6.6 % of trials). The average response time (RT) was 2208 ms (*SD* = 1919 ms). Trials that were more than 2.5 standard deviations above the mean RT (rounded to >7000 ms) were excluded from all analyses (1.9 % of trials). We did not further analyse the RTs and errors, as our predictions were all related to visual selection and the respective eye movement parameters.

6.1. First fixations: Training effects

To assess whether attentional priorities changed with experience, we analysed the probability of fixating on each of the differently coloured distractors as a function of trial repetition (see Fig. 4). The probability of fixating on each differently coloured distractor was computed by dividing the proportion of trials in which a given distractor was fixated (as the first item) by the number of distractors present on the trial (e.g., as there were two optimal-coloured stimuli, the proportion of first fixations on an optimal coloured stimulus was divided by two).

The proportion of first fixations on the target stimulus and the target-matching distractor did not differ from each other, $F(1,31) = 0.95$, $p = .337$, and did not interact with trial repetition, $F(9, 279) = 1.16$, $p = .321$. This shows that attention was not guided to the target item based on the contained response-related character, and thus the data were collapsed across the target and target-coloured distractor.

A 4 (Item Type: Target-Matching, Optimal, Relational, Salient) x 10 (Trial Repetition: T_1 to T_{10}) repeated measures analysis of variance (ANOVA) was computed over the probabilities of fixating on each differently coloured item. The results showed a main effect of item type, $F(3, 93) = 53.47$, $p < .001$, $\eta_p^2 = 0.63$, trial repetition, $F(9, 279) = 4.42$, $p < .001$, $\eta_p^2 = 0.13$, as well as a significant interaction, $F(27, 837) = 2.46$, $p < .001$, $\eta_p^2 = 0.07$.

To test our main hypotheses, we compared the proportion of first eye movements to each of the different distractors (relational, optimal and salient) to those of the combined target and target-similar distractor. To control for inflation of the α-error with repeated testing, we always report the FDR-corrected *p*-values (Benjamini & Hochberg, 1995), together with the Bayes factors (BF_{10}). The first results showed that both the relational and optimal items attracted more first eye movements than the target-coloured items (relational: $t(31) = 5.33$, $p < .001$, $BF_{10} = 2500.17$, optimal: $t(31) = 6.61$, $p < .001$, $BF_{10} = 7.24 * 10^4$), and did not differ from each other, $t(31) = 1.07$, $p = .294$, $BF_{10} = 0.32$. In turn, the target-coloured items received more first fixations than the saliently different item, $t(31) = 3.96$, $p < .001$, $BF_{10} = 71.41$.

To investigate possible linear trends that may reflect learning or adaptation effects, we next examined the slopes of first fixation locations separately for each of the differently coloured items across the ten trials. The linear trends for all four item types differed significantly from zero. The relational item had the steepest positive slope, $\beta = 0.33$, $t(31) = 4.38$, $p < .001$, followed by the optimal item, $\beta = 0.18$, $t(31) = 2.87$, $p = .007$, and the target-matching item, $\beta = 0.09$, $t(31) = 2.36$, $p = .025$, reflecting that selection increased for all of these items over the course of the block. The saliently different item showed a significant reduction in visual selection across blocks, $\beta = -0.11$, $t(31) = 2.08$, $p = .046$.

The finding that the relational and optimal distractors attracted the gaze more strongly than the target and target-similar items is in line with the relational account, but not with optimal tuning or feature-specific tuning accounts (compare Figs. 3 and 4). Thus, the first fixations in a trial show relational tuning, which became stronger across trials and remained dominant across all trials in the block.

6.2. First fixations: Initial guidance on T_1

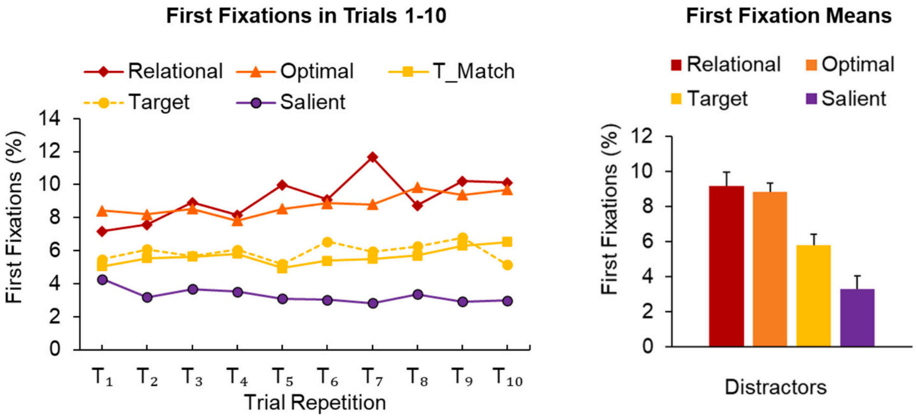
We next analysed fixations on the first trial of the block, when participants were unaware of the relative colour of the target (redder/yellower or greener/bluer), and only knew the exact target colour (orange or teal; see Fig. 4, T_1).

Paired-samples *t*-tests revealed that more first fixations on T_1 were made towards the relational distractor, $t(31) = 2.50$, $p = .036$, $BF_{10} = 2.71$ and the optimal distractor than the target, $t(31) = 4.34$, $p < .001$, $BF_{10} = 185.7$, whereby optimal and relational distractors did not differ significantly from each other, $t(31) = 1.59$, $p = .147$, $BF_{10} = 0.58$. The salient item did not differ significantly from the target-coloured item, $t(31) = 0.88$, $p = .388$, $BF_{10} = 0.27$ on the first trial.

This results pattern most closely matches the predictions of the relational account (see Fig. 3). The finding that relational search was conducted on trial 1 shows that experience or prior knowledge is not necessary for relational search, and supports the hypothesis that the visual system can quickly assess the dominant colour in the visual field and tune attention to the relative colour of the target.

6.3. Fixation progression

The results from the first fixations above suggest that search was initially (on the first trial) relational and that this relational guidance



First Fixations in Trials 1-10

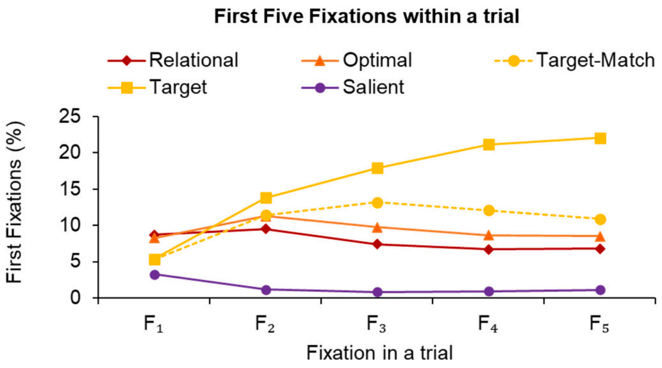
First Fixation Means

Fig. 4. Left: The proportion of first fixations directed towards the different distractors (relational, optimal, target-matching, or salient) or the target, depicted as a function of trial repetition (Trial 1–10). The relational and optimal distractors attracted the highest proportions of first fixations and showed a linear increase over the ten trials. The target and the target-matching distractor received significantly less fixations than the relational or optimal distractors, but more first fixations than the saliently different item, and also displayed a linear increase in fixations. The salient item was selected least frequently and showed a linear decrease in first fixations. Right: Mean proportion of first fixations on each item, averaged over trial repetition. Error bars depict +1 SEM.

was maintained and increased across repeated trials. We next assessed whether the relational account accurately predicts the fixation progression *within a single* trial, with attention first selecting the most extreme relatively matching item (e.g., reddest), followed by the optimal (e.g., next-reddest) item, before honing in on the target-coloured items. To that aim, we analysed the first five fixations in each trial (see Fig. 5). On average participants made 4.58 fixations per trial. We included all trials in this data set, including those where the task was completed within the first five fixations (which resulted in 1.7 % missing data for F_2, F_3: 7.4 %, F_4: 17.5 %, and F_5: 30.6 %). We computed all *t*-tests separately for the target and target-similar distractor now, using FDR correction to adjust the *p*-values and also reported the results of Bayesian statistics.

The first fixation (F_1), now collapsed over trial repetition showed that both relational and optimal stimuli attracted a higher proportion of first fixations than both the target and the target-matching distractor (all $ts > 4.7$, $ps < 0.001$, $\mathrm{BF}_{10} > 511.08$). For F_2, results began to deviate from relational guidance. The target stimulus was now more likely to be fixated than any other item, all $ts > 2.76$, $ps \leq 0.011$, $\mathrm{BF}_{10} > 4.59$. The target-matching distractors were now equally frequently fixated as the optimal and relational distractors, both $ts < 1.7$, $ps > 0.12$, $\mathrm{BF}_{10} < 0.64$. From F_3 onwards fixation patterns remained consistent: Now the target-matching distractors were more likely to be fixated than the optimal distractor (all $ts > 3.5$, $ps \leq 0.001$, $\mathrm{BF}_{10} > 25.52$), and the optimal distractor was more likely to be fixated than the relational distractor (all $ts > 5.1$, $ps < 0.001$, $\mathrm{BF}_{10} > 1361.58$). Finally, fixations on the salient item were highest on F_1 than at any other point in the trial (all $ts > 5.9$, $ps < 0.001$, $\mathrm{BF}_{10} > 14{,}417.06$) and were significantly lower than all other stimuli of interest across all fixations (all $ts > 3.1$, $ps \leq 0.004$, $\mathrm{BF}_{10} > 10.67$).

In sum, the results are broadly in line with the hypothesis that participants will first select the relationally maximal item (e.g., reddest), followed by the optimal (e.g., next-reddest) item, and the target-coloured items. Another way of describing these results is that attention was always initially guided to the relatively matching items, with feature-specific tuning developing only after the first (few) fixation(s).



First Five Fixations within a trial

Fig. 5. The first five fixations (F_1 to F_5) in a trial show how search progressed during a trial. The first fixation (F_1) displayed relational search, with the relational and optimal distractors were selected more frequently than the target and target-matching distractor. The second fixation (F_2) already showed a shift to more feature-specific guidance: The target was now the most likely to be fixated, followed by the target-matching and optimal distractors, with fewer fixations on the relational distractor. With the third fixation (F_3), the target-matching distractor was now more likely to be fixated than optimal or relational distractors. From this point onwards, fixation proportions seem to be dictated by target-feature similarity. The salient item was most frequently fixated on F_1 and then was hardly selected anymore in subsequent fixations.

6.4. *First saccade latencies*

We next analysed the saccade latencies, that is, the time from the onset of the search display to the onset of the first saccade in a trial, to assess the time-course of attentional deployment using the first saccade in a trial. This analysis was collapsed over trials, with the first trial being excluded (due to longer saccadic latencies).[3] Fig. 6 shows the distributions of saccadic latencies separately for each item type (i.e., the proportion of trials where saccadic latencies ranged from 125 ms – 150 ms, 150–175 ms, etc.).

To analyse the data statistically, we fed the average saccade latencies of the first saccades in a trial into a one-way (Item Type: Target, Optimal, Relational, Salient) repeated-measures ANOVA. The results showed a significant effect of item type $F(3,93) = 53.01$, $p < .001$, $\eta_p^2 = 0.63$. Saccade latencies were shortest for the salient distractor ($M = 218$ ms), significantly shorter than for both the relational ($M = 237$ ms) and optimal distractors ($M = 241$ ms), all $ts > 8.18$, $ps < 0.001$, $\mathrm{BF}_{10} > 3.98 \times 10^6$. Saccade latencies to the relational and optimal distractors were shorter than to target-matching items ($M = 252$ ms), all $ts > 4.35$, $ps < 0.001$, $\mathrm{BF}_{10} > 191.72$, while relational and optimal distractors did not differ, $t(31) = 1.26$, $p = .218$, $\mathrm{BF}_{10} = 0.39$.

---

[3] Saccade latencies on T_1 were longer than in all other trial repetitions (all *ps* < 0.001), but there was no repetition by item type interaction ($p = .650$), or any linear trends for the individual items after excluding the first trial (*ps* > 0.235).

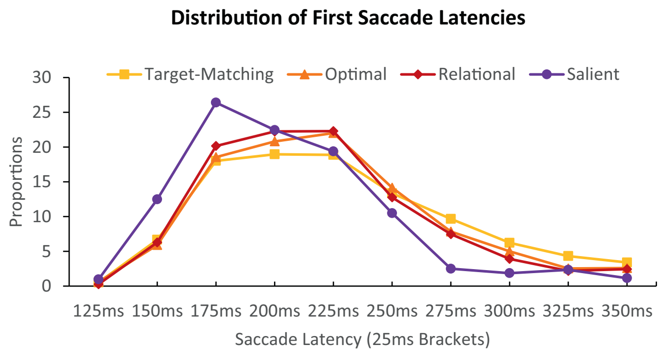


Fig. 6. The distribution of saccade latencies, which was derived by sorting saccades to each stimulus type into 25 ms latency bins and depicting the proportions of trials within each bin separately for each item type. The results showed a higher proportion of short-latency saccades to the saliently different distractor than to the other distractors. Moreover, there was a higher proportion of saccades with long latencies to target-matching items (pooled over target and target-coloured distractor), producing differences between the target-coloured items and the other distractors in the tail end of the distribution.

These results show that saccades to the salient item were initiated significantly earlier than to optimal and relational distractors, which in turn were selected earlier than the target-similar items. Thus, selection was initiated earlier the more dissimilar the distractors were from the target item. By contrast, the rather long saccade latencies for the target and target-matching distractors were due to a higher proportion of saccades with longer latencies.

7. Distribution of first saccade latencies

7.1. Distractor dwell times

As in previous studies, we also examined the mean dwell times on each of the distractors, measured as the time spent fixating on each of the different item types. Target fixations were excluded from this analysis, and results were collapsed over trial repetition. The one-way (Distractor Type: Target-Matching, Optimal, Relational, Salient) repeated-measures ANOVA revealed significant differences between the item types, $F(3,93) = 40.50$, $p < .001$, $\eta_p^2 = 0.57$. Dwell times were shortest for the salient item ($M = 152$ ms), significantly shorter than for the relational distractor (166 ms), $t(31) = 6.26$, $p < .001$, $\mathrm{BF}_{10} = 2.90 * 10^4$. Dwell times on the optimal ($M = 171$ ms) and target-matching distractors ($M = 173$ ms) were longest, significantly longer than on the relational distractor (optimal distractor: $t(31) = 5.74$, $p < .001$, $\mathrm{BF}_{10} = 7312.78$; target matching distractors: $t(31) = 3.21$, $p = .004$, $\mathrm{BF}_{10} = 11.99$), and did not differ from each other, $t(31) = 0.92$, $p = .363$, $\mathrm{BF}_{10} = 0.28$.

In sum, dwell times increased with similarity to the target colour, indicating that longer dwell times were needed to identify the items as distractors as they became more similar to the target.

7.2. General discussion

The current study yielded several important results that significantly expand our understanding of attentional tuning. This was the first study to examine attentional tuning functions in the absence of target contextual knowledge and in displays containing numerous items, when the target never had a superlative relative feature (e.g., never the reddest item). Still, relatively matching distractors reliably attracted the first eye movement, including on the *first* trial in a mini-block, when participants only knew the colour of the target but not its relative contextual colour. Of note, we failed to find any differences between relational and optimal distractors on T_1 fixations. This may indicate that the colours were too similar to each other to be pre-attentively discriminable. Still, the finding that the relatively more extreme distractors were selected first shows that contextual information can be quickly extracted from a visual scene and used to guide attention.

Previous studies have already shown that information about the properties of a visual scene can be rapidly extracted, as reflected in our ability to accurately estimate the average feature of multiple objects in a visual scene (feature averaging, e.g., Chong & Treisman, 2005a, 2005b; Choo & Franconeri, 2010; Joo, Shin, Chong, & Blake, 2009), and our ability to rapidly process the gist of a visual scene (e.g., Oliva & Torralba, 2006; see also Thorpe, Fize, & Marlot, 1996). The present study extends on this research by showing that contextual information can be rapidly extracted and used to bias attention to the relative features of the target, prior to the first eye movement. While the exact mechanism that allows us to extract this information is currently unclear, the present study clearly shows that relational guidance does not require learning of the context or knowledge of the relative target feature, but can be executed with the first glance at a novel scene.

The current results also revealed that relational guidance is more ubiquitous than previously thought. In particular, we found relational guidance even when the target never had a relationally maximal value – that is, when it was never the reddest, greenest, bluest or yellowest item in the display. In every trial there were relationally better-matching distractors (one relatively matching distractor and two optimal distractors), which frustrated selecting the target with the first eye movement in relational search. Yet, we reliably found that participants searched relationally. This indicates that the visual system is indeed sensitive to the dominant feature in the display, and initiates relational search as soon as the target differs from *most* of the display items in a single direction. Furthermore, participants continued to search relationally over as many as ten successive trials, even though relational guidance reliably resulted in selecting one of the more extreme distractors first. This means that relational guidance is more readily applied and more persistent than previously thought (e.g., Becker, Valuch, and Ansorge, 2014).

The results of the present study also provided new insights into the progression of visual search. Selection seemed to largely proceed in the way as laid out in the relational account, with the most extreme, relatively matching item being selected first (e.g., reddest), followed by the next most extreme, relatively matching item (i.e., next reddest item), and so forth, until the target is selected. As shown in Fig. 3, the first saccade in a trial was most likely to select relatively matching items, whereas later saccades were more likely to select target-matching items. As the probability of selecting target-matching items increased, the probability for selecting extreme items declined (especially after the second fixation).

This sequence of eye movements best matches the predictions of the relational account. The results only deviate from the predictions in that the relatively matching and optimal distractors mostly had equally high selection rates that did not differ from each other. The failure to find more frequent selection of the relational than the optimal distractor is probably due to the fact that the relational and optimal colours were not distinct enough to allow rapid discrimination and identifying the truly maximal colours (i.e., reddest, bluest, yellowest, greenest colours). In line with this possibility, selection of target-matching items occurred slightly earlier than expected, with the 2nd fixation, indicating that relational and optimal colours were grouped together and rejected as a group after the 1st fixation on either item (e.g., Duncan & Humphreys, 1989). While this explanation warrants further research, the data clearly showed that the progression of visual search is a guided process that progresses from relatively matching, extreme items to target-matching items that share the target's features.

Other theories have proposed a memory-based, 'inhibitory tagging' or 'inhibition of return' mechanism to explain how search proceeds to the next item (e.g., Klein, 1988, 2000), which has been implemented into most models (e.g., Wolfe, 2021). Of note, inhibitory tagging could be either location-based (i.e., inhibiting selected nontarget locations during the trial), or feature-based (i.e., inhibiting a non-target or distractor colour after selection; e.g., Bichot & Schall, 2002; Braithwaite, Humphreys, & Hulleman, 2005), in which case it would prevent re-fixating the same colours.

Inhibitory tagging implemented as location-based inhibition cannot explain the current results because selection of the target and target-matching distractors commenced earlier than would be expected on the basis of location-based inhibition. Specifically, as there were three relatively matching or optimal distractors in the display plus three target-matching items, the data should have shown continued selection and inhibition of the locations containing candidate target items before honing in on the target. By contrast, we observed a strong decline in selecting distractors after the 2nd fixation, which corresponds to the number of different distractor colours (rather than locations). Also, selection of the target and target-matching items showed a steep increase after only one fixation, which seems too early for location-based inhibition. Thus, if inhibitory tagging is invoked to explain the current results, we have to assume that the colour(s) rather than the locations of the more extreme distractors were inhibited.

Alternatively, it is possible that attention was tuned to the exact feature value of the target after the first initial selections on relatively maximal items, swiftly switching from originally relational search to a narrower, more feature-specific target template (e.g., Duncan & Humphreys, 1989). In line with this possibility, previous studies have shown that we can bias attention to the specific feature value of the target when relational search does not allow efficient selection of the target (e.g., Becker, Valuch, & Ansorge, 2014a, 2014b; Harris, Remington, & Becker, 2013; Schönhammer, Grubert, Kerzel, & Becker, 2016). Hence, it is possible that search is initially driven by a relational target template, which narrows to a more feature-specific search template once a relatively matching distractor has been selected (see also Grössle et al., 2023).

While the current study cannot distinguish between these two different explanations, it is important to note that inhibition or adaptations of the target template that occurred within a trial had transient, short-lived effects. Usually, an attentional bias for either relative features or specific feature values will automatically carry over to the next trial to influence selection (i.e., intertrial priming; Maljkovic & Nakayama, 1994; see also Becker, Valuch, & Ansorge, 2014a). Similarly, inhibition of a distractor feature will usually also carry over to the next trial to influence selection, though its effects are usually weaker than effects of target guidance (e.g., Chang & Egeth, 2019; Hamblin-Frohman et al., 2022). However, in the present study there was no evidence that inhibition or adaptations to the search template carried over to the next trial: Relational effects became *stronger* over the course of the block rather than weaker (see Fig. 4). Thus, adaptations to search that were made to allow progression to the next item within a trial were short-lived and did not affect search on the next trial(s).

By contrast, we also found evidence of inhibition that carried over to the next trial to influence selection: The salient distractor only attracted attention on the first trial in a mini-block, and was successfully ignored after the first trial (see Fig. 4). This result is in line with previous inhibition studies (e.g., Gaspelin & Luck, 2018; Hamblin-Frohman et al., 2022), and extends on these findings by showing that the salient distractor can also be successfully inhibited or suppressed in larger displays comprising 36 items (see also Hamblin-Frohman, Pratt & Becker, in press and Ramgir & Lamy, 2023, for similar findings). These findings are at odds with the results of Wang and Theeuwes (2020), who only found inhibition in 4-item displays but not in larger displays containing 6 or 10 items. It is possible that they failed to find inhibition in larger displays because their displays contained only a few different shapes and colours, which may have promoted a singleton detection strategy in search for the target, which prevented inhibition of salient items (e.g., Bacon & Egeth, 1994; Hamblin-Frohman, Pratt, & Becker, 2025; but see Becker, Martin, & Hamblin-Frohman, 2019). While this explanation requires further investigation, our results clearly show that irrelevant salient distractors can be suppressed even in large display sizes if the conditions encourage tuning attention to the (relative) target feature. With this, the current study confirms previous results with sparser displays (Martin & Becker, 2018; York & Becker, 2020; see also Hamblin-Frohman, Pratt & Becker, in press), as well as previous studies showing distractor suppression when saliency was manipulated or measured in a different manner (e.g., Stilwell, Adams, & Egeth, 2023).

It should be noted that inhibition of the salient distractor was also fairly strong, completely suppressing capture by the salient item. Averaged across all trials, salient items actually attracted *fewer* first eye movements (3.3 %) than the average, non-salient non-targets (4.3 %).[4] Given the low proportion of first fixations on the salient item, we cannot claim that the salient item *overall* attracted attention. Even in the densely populated displays, top-down tuning to the (relative) target feature modulated selection rates of the differently coloured distractors more strongly than bottom-up feature contrasts, indicating that top-down processes dominate attentional guidance over bottom-up saliency. These results are in line with previous studies showing that top-down processes dominate attentional guidance (e.g., Becker et al., 2017) and invalidate bottom-up saliency views which claim that bottom-up saliency is the most important factor in attentional guidance (e.g., Theeuwes, 2004; Wang & Theeuwes, 2020).

These results also rule out an alternative explanation of the results. Proponents of feature-based theories (e.g., Guided Search 2.0; Wolfe, 1994) may argue that the results could be due to a combination of top-down, feature-specific tuning to the target and bottom-up processes, as the more extreme distractors (optimal and relatively matching distractors) may have been similar enough to the target to be subject to (broad) top-down tuning and were simultaneously more salient than the target, leading to higher selection rates. Contrary to this contention, we found a very different results pattern for the salient distractor compared to the relatively matching distractors, in two respects: First, the trial-by-trial analysis showed an increase in capture by relatively matching and target-similar distractors over trials but a *de*crease for the salient distractor (see Fig. 3). Second, the time-course analysis showed shorter latencies of saccades to the salient distractor than for all other distractors (see Fig. 5). These findings suggest that the more extreme (and more salient) relatively matching and optimal distractors attracted attention and the gaze in the same manner as the target-similar (less salient) distractor, viz., due to top-down tuning and not bottom-up saliency.

These findings are in line with previous work (e.g., Becker et al., 2013) and extend on it, by showing that attention can also be tuned to relative features even when the target is never the relatively reddest, bluest, greenest or yellowest item in the display and only differs from the majority of non-targets in a linear fashion. Although the conditions promoted tuning to the specific feature value of the target by persistently displaying relatively matching distractors, participants consistently engaged in relational search and also did not adjust search towards adopting a more feature-specific bias for the target. To the contrary: As shown in Fig. 3, linear trends for first eye movements increased for capture by the relational and the optimal distractors compared to the target-matching items, suggesting that the relational effect became *stronger* over the course of the block rather than weaker.

Collectively, these results show that participants can be quite reluctant to tune attention to the specific feature value of the target. We can only speculate why this may be the case. Previous studies have

---

[4] To compare the proportion of first fixations on the salient item to the proportion of fixations on the non-salient non-target items, we omitted the corner positions and central positions that never contained a distractor, and compared the resulting proportions using a paired, two-tailed *t*-test. The results revealed anecdotal evidence that the non-salient non-target items were selected more frequently (4.3 %) than the salient item (3.3 %, with the first eye movement), $t(31) = 2.22$, $p = .034$, $\mathrm{BF}_{10} = 1.61$.

shown that feature-specific tuning results in delays in selecting the target, as well as longer dwell times on target-similar distractors, compared to tuning to relations (in identical displays; Becker, Valuch, & Ansorge, 2014a, 2014b; Martin & Becker, 2018). This suggests that tuning attention to exact feature values may cause delays in selection as well as decision-making (about whether the selected item is the target), rendering feature-specific search less efficient than relational search. While this question warrants further investigation, the results clearly show that relational search does not require the target to have superlative feature on a portion of trials and can be observed once the target differs from the majority of non-targets in a relative feature (e.g., redder).

8. Conclusion

In summary, the present study showed that tuning to relative features occurs in more densely populated, 36-item displays; in the absence of knowledge of the relative feature of the target, and in the absence of any training. This demonstrates the existence of a mechanism that can extract the dominant feature in a visual scene and using that information to tune attention to the relative feature of a target. Moreover, we found relational search in displays that contained multiple different colours, and even though the target never had a relatively maximal feature, which indicates that relational search is likely a default search mode also in everyday situations (see also Yu, Zhou, Becker, Boettcher and Geng, 2023). In addition, the results showed solid evidence of inhibition of a salient distractor after the first presentation of it, which prevented further selection of it, demonstrating that task-oriented processes can dominate bottom-up saliency in larger and more ecologically valid displays. Finally, we found evidence that search progresses from the most relatively extreme item to the next-extreme relatively matching item until the target is found, demonstrating the existence of a feature-specific mechanism that guides fixations and attention to the most likely candidate target within a given trial. These results confirm several important tenets of the relational account and demonstrate the importance of context-dependent information in the guidance of attention (Yu et al., 2023).

This research was supported by Australian Research Council (ARC) grants DP210103430 and DP240102774 to SIB. There are no conflicts of interest to report. This study was not pre-registered. All materials are available on request. Raw and averaged data are available at: https://osf.io/74z2s

CRediT authorship contribution statement

Stefanie I. Becker: Writing – review & editing, Supervision, Software, Resources, Project administration, Funding acquisition, Conceptualization. **Zachary Hamblin-Frohman:** Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Koralalage Don Raveen Amarasekera:** Data curation.

Data availability

We have included a link to the full data set in the paper

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2025.106132.

References

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science, 28*(11), 1547–1562. https://doi.org/10.1177/0956797617723724

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*, 485–496. https://doi.org/10.3758/BF03205306

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision Research, 36*(10), 1439–1465. https://doi.org/10.1016/0042-6989(95)00207-3

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

Becker, S. I. (2013a). Why you cannot map attention: A relational theory of attention and eye movements. *Australian Psychologist, 48*, 389–398. https://doi.org/10.1111/ap.12028

Becker, S. I. (2013b). Simply shapely: Relative, not absolute shapes are primed in pop-out search. *Attention, Perception, and Psychophysics, 75*(5), 845–861. https://doi.org/10.3758/S13414-013-0433-1/FIGURES/4

Becker, S. I. (2018). Reply to Theeuwes: Fast, feature-based top-down effects, but saliency may be slow. *Journal of Cognition, 1*(28), 1–3. https://doi.org/10.5334/joc.23

Becker, S. I., Folk, C. L., & Remington, R. W. (2013). Attentional capture does not depend on feature similarity, but on target-nontarget relations. *Psychological Science, 24*(5), 634–647. https://doi.org/10.1177/0956797612458528

Becker, S. I., Grubert, A. K., Ansorge, U., & Horstmann, G. (2023). Which processes dominate visual search: Bottom-up feature contrast, top-down tuning or trial history? *Cognition, 236*, Article 105420. https://doi.org/10.1016/j.cognition.2023.105420

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*(1), 264–291. https://doi.org/10.1037/a0033489

Becker, S. I., Harris, A. M., York, A., & Choi, J. (2017). Conjunction search is relational: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 43*(10), 1828. https://doi.org/10.1037/XHP0000371

Becker, S. I., Martin, A., & Hamblin-Frohman, Z. (2019). Target templates in singleton search vs. feature-based search modes. *Visual Cognition, 27*, 502–517. https://doi.org/10.1080/13506285.2019.1676352

Becker, S. I., Retell, J. D., & Wolfe, J. M. (2023). Mirror blindness: Our failure to recognise the target in search for mirror-reversed shapes. *Attention, Perception & Psychophysics, 85*, 418–437. https://doi.org/10.1016/j.actpsy.2014.10.002

Becker, S. I., Valuch, C., & Ansorge, U. (2014a). Color priming in pop-out search depends on the relative color of the target. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00289

Becker, S. I., Valuch, C., & Ansorge, U. (2014b). Color priming in pop-out search depends on the relative color of the target. *Frontiers in Psychology, 5*(289), 1–11. https://doi.org/10.3389/fpsyg.2014.00289

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple hypothesis testing. *Journal of the Royal Statistical Society B, 57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bichot, N. P., & Schall, J. D. (2002). Priming in macaque frontal cortex during Popout visual search: Feature-based facilitation and location-based inhibition of return. *Journal of Neuroscience, 22*(11), 4675–4685. https://doi.org/10.1523/JNEUROSCI.22-11-04675.2002

Braithwaite, J. J., Humphreys, G. W., & Hulleman, J. (2005). Color-based grouping and inhibition in visual search: Evidence from a probe detection analysis of preview search. *Perception & Psychophysics, 67*(1), 81–101. https://doi.org/10.3758/BF03195014

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

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484–1525. https://doi.org/10.1016/J.VISRES.2011.04.012

Chang, S., & Egeth, H. E. (2019). Enhancement and suppression flexibly guide attention. *Psychological Science, 30*(12), 1724–1732. https://doi.org/10.1177/0956797619878813

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

Chong, S. C., & Treisman, A. (2005b). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891–900. https://doi.org/10.1016/j.visres.2004.10.004

Choo, H., & Franconeri, S. L. (2010). Objects with reduced visibility still contribute to size averaging. *Attention, Perception, & Psychophysics, 72*(1), 86–99. https://doi.org/10.3758/APP.72.1.86

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837. https://doi.org/10.1016/0042-6989(95)00294-4

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

D'Zmura, M. (1991). Color in visual search. *Vision Research, 31*(6), 951–966. https://doi.org/10.1016/0042-6989(91)90203-h

Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences, 22*(1), 79–92. https://doi.org/10.1016/J.TICS.2017.11.001

Geng, J. J., di Quattro, N. E., & Helm, J. (2017). Distractor probability changes the shape of the attentional template. *Journal of Experimental Psychology: Human Perception and Performance, 43*(12), 1993. https://doi.org/10.1037/XHP0000430

Grössle, I. M., Schubö, A., & Tünnermann, J. (2023). Testing a relational account of search templates in visual foraging. *Scientific Reports, 13*(1), 12541. https://doi.org/10.1038/s41598-023-38362-9

Hamblin-Frohman, Z., & Becker, S. I. (2021). The attentional template in high and low similarity search: Optimal tuning or tuning to relations? *Cognition, 212*, Article 104732. https://doi.org/10.1016/j.cognition.2021.104732

Hamblin-Frohman, Z., Chang, S., Egeth, H., & Becker, S. I. (2022). Eye movements reveal the contributions of early and late processes of enhancement and suppression to the guidance of visual search. *Attention, Perception, and Psychophysics, 84*(6), 1913–1924. https://doi.org/10.3758/S13414-022-02536-W/FIGURES/4

Hamblin-Frohman, Z., Pratt, J., & Becker, S. I. (2025). *Inhibition in large set sizes depends on search mode, not salience.* Attention, Perception & Psychophysics.

Harris, A. M., Remington, R. W., & Becker, S. I. (2013). Feature specificity in attentional capture by size and colour. *Journal of Vision, 13*(12), 1–15. https://doi.org/10.1167/13.3.12

Horstmann, G., & Becker, S. I. (2020). More efficient search for happy targets may not indicate guidance, but rather faster distractor rejection: Evidence from eye movements and fixations. *Emotion, 20*, 206–216. https://doi.org/10.1037/emo0000536

Huang, L., & Pashler, H. (2005). Expectation and repetition effects in searching for featural singletons in very brief displays. *Perception & Psychophysics, 67*, 150–157. https://doi.org/10.3758/BF03195018

Joo, S. J., Shin, K., Chong, S. C., & Blake, R. (2009). On the nature of the stimulus information necessary for estimating mean size of visual arrays. *Journal of Vision, 9*(9), 1–12. https://doi.org/10.1167/9.9.7

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

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*(4), 138–147. https://doi.org/10.1016/S1364-6613(00)01452-2

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657–672. https://doi.org/10.3758/BF03209251/METRICS

Martin, A., & Becker, S. I. (2018). How feature relationships influence attention and awareness: Evidence from eye movements and EEG. *Journal of Experimental Psychology: Human Perception and Performance, 44*(12), 1865. https://doi.org/10.1037/XHP0000574

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

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

Oliva, A., & Torralba, A. (2006). Chapter 2 building the gist of a scene: The role of global image features in recognition. *Progr. Brain Res.*, 23–36. https://doi.org/10.1016/S0079-6123(06)55002-2

Oriet, C., & Brand, J. (2013). Size averaging of irrelevant stimuli cannot be prevented. *Vision Research, 79*, 8–16. https://doi.org/10.1016/j.visres.2012.12.004

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. https://doi.org/10.1016/J.JNEUMETH.2006.11.017

Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede involuntary eye movements. *Perception & Psychophysics, 66*(3), 398–405. https://doi.org/10.3758/BF03194888

Ramgir, A., & Lamy, D. (2023). Distractor's salience does not determine feature suppression: A commentary on Wang and Theeuwes (2020). *Journal of Experimental Psychology: Human Perception and Performance, 49*(6), 852–861. https://doi.org/10.1037/xhp0001119

Schönhammer, J. G., Grubert, A., Kerzel, D., & Becker, S. I. (2016). Attentional guidance by relative features: Behavioral and electrophysiological evidence. *Psychophysiology, 53*(7), 1074–1083. https://doi.org/10.1111/psyp.12645

Scolari, M., Byers, A., & Serences, J. T. (2012). Optimal deployment of attentional gain during fine discriminations. *Journal of Neuroscience, 32*(22), 7723–7733. https://doi.org/10.1523/JNEUROSCI.5558-11.2012

Stilwell, B. T., Adams, O. J., Egeth, H. E., et al. (2023). The role of salience in the suppression of distracting stimuli. *Psychonomic Bulletin and Review, 30*, 2262–2271. https://doi.org/10.3758/s13423-023-02302-5

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*(1), 65–70. https://doi.org/10.3758/BF03206462

Theeuwes, J. (2013). Feature-based attention: It is all bottom-up priming. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1628). https://doi.org/10.1098/RSTB.2013.0055

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522. https://doi.org/10.1038/381520a0

Wang, B., & Theeuwes, J. (2020). Salience determines attentional orienting in visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 46*(10), 1051–1057. https://doi.org/10.1037/xhp0000796

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

Wolfe, J. M. (2020). Visual search: How do we find what we are looking for? *Annual Review of Vision Science, 6*(1), 539–562. https://doi.org/10.1146/annurev-vision-091718-015048

Wolfe, J. M. (2021). Guided search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review 2021 28:4, 28*(4), 1060–1092. https://doi.org/10.3758/S13423-020-01859-9

Wyble, B., Folk, C. L., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 861–871. https://doi.org/10.1037/a0030517

York, A., & Becker, S. I. (2020). Top-down modulation of gaze capture: Feature similarity, optimal tuning, or tuning to relative features? *Journal of Vision, 20*(4), 1–16. https://doi.org/10.1167/JOV.20.4.6

Yu, X., Hanks, T. D., & Geng, J. J. (2022). Attentional guidance and match decisions rely on different template information during visual search. *Psychological Science, 33*(1), 105–120. https://doi.org/10.1177/09567976211032225/ASSET/IMAGES/LARGE/10.1177_09567976211032225-FIG 2.JPEG

Yu, X., Zhou, Z., Becker, S. I., Boettcher, S. E. P., & Geng, J. (2023). Good enough attentional guidance. *Trends in Cognitive Sciences, 37*, 391–403.

Zelinksy, G., & Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 244–262. https://doi.org/10.1037//0096-1523.23.1.244. doi:10.3758/s13414-022-02641-w