Which processes dominate visual search: Bottom-up feature contrast, top-down tuning or trial history?

Stefanie I. Becker[1] , Anna Grubert[2], Gernot Horstmann[3], Ulrich Ansorge[4]

1 The University of Queensland, Brisbane, Australia

2 Durham University, United Kingdom

3 University of Bielefeld, Germany

4 The University of Vienna, Austria

Abstract

Previous research has identified three mechanisms that guide visual attention: bottom-up feature contrasts, top-down tuning, and the trial history (e.g., priming effects). However, only few studies have simultaneously examined all three mechanisms. Hence, it is currently unclear how they interact or which mechanisms dominate over others. With respect to local feature contrasts, it has been claimed that a pop-out target can only be selected immediately in dense displays when the target has a high local feature contrast, but not when the displays are sparse, which leads to an inverse set-size effect. The present study critically evaluated this view by systematically varying local feature contrasts (i.e., set size), top-down knowledge, and the trial history in pop-out search. We used eye tracking to distinguish between early selection and later identification-related processes. The results revealed that early visual selection was mainly dominated by top-down knowledge and the trial history: When attention was biased to the target feature, either by valid pre-cueing (top-down) or automatic priming, the target could be localised immediately, regardless of display density. Bottom-up feature contrasts only modulated selection when the target was unknown and attention was biased to the non-targets. We also replicated the often-reported finding of reliable feature contrast effects in the mean RTs, but showed that these were due to later, target identification processes (e.g., in the target dwell times). Thus, contrary to the prevalent view, bottom-up feature contrasts in dense displays do not seem to directly guide attention, but only facilitate nontarget rejection, probably by facilitating nontarget grouping.

1. Introduction

We cannot consciously perceive all the information contained in a visual scene at once (e.g., Helmholtz, 1867; James, 1890; Simons & Levin, 1997). Selective attention comprises a set of mechanisms that allows us to prioritise certain objects over others. Previous research has identified at least three main mechanisms that can drive visual attention (Awh, Belopolsky, & Theeuwes, 2012): (1) a bottom-up, saliency based mechanism that drives attention to the items with the highest (local) feature contrasts (e.g., Itti & Koch, 2000; Itti, Koch, & Niebur, 1998; Li, 2002; Theeuwes, 1992, Theeuwes, 2004; Wolfe, 1994), (2) a top-down, feature-based mechanism that allows us to selectively attend to task-relevant items (e.g., Folk & Remington, 1998; Wolfe, 1994), and (3) a memory-based system that drives attention according to the search history, by *priming* attention to select items that are similar to previously selected items (e.g., Maljkovic & Nakayama, 1994; McPeek, Maljkovic, & Nakayama, 1999).

According to the most prominent models of visual attention, feature contrast affects visual selection automatically and independently of the intentions and goals of the observer, probably due to hard-wired properties of the visual system (e.g., the anatomy and response characteristics of visual neurons, e.g., Itti & Koch, 2000; Martinez-Trujillo & Treue, 2004; Navalpakkam and Itti, 2006, Navalpakkam and Itti, 2007; Reynolds & Desimone, 2003; Sprague, Itthipuripat, Vo, & Serences, 2018; Van Zoest, Donk, & Theeuwes, 2004; White, Kan, Levy, Itti, & Munoz, 2017; Wolfe, 1994, Wolfe, 2021). By contrast, top-down tuning and priming are part of a more dynamic 'feature weighting' system that adjusts the gain of specific features according to current task demands or the search history (e.g., by increasing the response of neurons that respond to specific feature values, e.g., Koch & Ullman, 1985; Maunsell & Treue, 2006; Navalpakkam & Itti, 2007; Spitzer, Desimone, & Moran, 1988; Wolfe, 1994). Thereby, feature priming effects are thought to occur automatically and without assistance from goal-driven processes, whereas top-down tuning describes changes in feature weighting that depend on the intentions and goals of the observers. Thus, these three mechanisms are thought to operate independently of each other.

Bottom-up feature contrasts play an important role in target selection: An important pre-requisite for efficient target detection is that the search target differs from the non-targets by a ‘just noticeable difference’ (*jnd*; e.g., Nagy & Sanchez, 1990). If the target is too similar to the non-targets, the target cannot be selected as the first item in the search display, instead requiring an effortful, piecemeal search (e.g., Duncan & Humphreys, 1989). However, this finding is consistent with two different mechanisms. According to Treisman and Gelade (1980), feature contrast should only play a role insofar as it promotes the emergence of elementary features that can be represented on separate feature maps. So increasing the feature contrast of an item beyond that point (e.g., beyond two pre-attentive jnd's) should have no effects. This exemplifies the view that feature contrast imposes a bottom-up limitation on selection, or that it is a necessary pre-condition for efficient search (but does not modulate search beyond that point). This contrasts with the view of Theeuwes (1992) that bottom-up feature contrasts modulate search continuously on a larger scale, beyond boundary conditions set by pre-attentive jnd's. Currently, it is unclear which of the two hypotheses or models is more accurate.

A widely held belief is that bottom-up mechanisms precede and, in this sense, determine top-down and priming mechanisms in attentional guidance (e.g., Theeuwes, 1992). However, this conclusion is by far not certain. To date, only very few studies have examined all three mechanisms in parallel (e.g., Leonard & Egeth, 2008). Several studies tested whether top-down tuning can override feature contrast effects by examining the effects of a salient task-irrelevant distractor on visual search performance (e.g., Becker, 2007; Becker, Lewis, & Axtens, 2017; Eimer, Kiss, Press, & Sauter, 2009; Theeuwes, 2004; Wang & Theeuwes, 2020). The results of these studies are still controversially discussed, and there seems to be little prospect for a resolution in the near future (Luck, Gaspelin, Folk, Remington, & Theeuwes, 2021; for a meta-analysis, see Büsel, Voracek, & Ansorge, 2020).[1]

Another set of studies investigated whether *local* feature contrasts can impose bottom-up limitations on pop-out search, by comparing search performance in dense displays where the target had a high local feature contrast versus sparse displays, in which the target had a low local feature contrast (e.g., Bravo & Nakayama, 1992; Meeter & Olivers, 2006; Rangelov, Müller, & Zehetleitner, 2017; see also Nakayama & Martini, 2011). The target in these studies was always a high-contrast singleton target (e.g., a red item among all-green nontargets; always present), and participants had to respond to an additional target feature. The local feature contrast of the items was manipulated by varying the number of items (set size) in the display. The first study of that kind (Bravo & Nakayama, 1992) found that increases in local feature contrasts (increased set size) benefited search when the target and nontarget colours randomly switched across trials, resulting in an inverse set size effect (i.e., faster search in dense displays). However, when the colours remained constant, increasing the local feature contrast or set size did not benefit search (as reflected in a flat set size function).

Bravo and Nakayama (1992) concluded that local feature contrasts can only modulate attention when the target is unknown – that is, when it is not possible to tune attention to the target colour in a top-down controlled fashion. According to this 'top-down' view, bottom-up feature contrasts can only modulate visual selection in the absence of top-down tuning, indicating that top-down tuning dominates over bottom-up feature contrasts in the guidance of attention. However, Bravo and Nakayama (1992) did not examine feature priming effects in pop-out search, as these were discovered only later (by Maljkovic & Nakayama, 1994).

A subsequent study by Rangelov et al. (2017) examined feature priming effects in high versus low contrast displays in a similar search task with 3, 12, or 36 items, and found results that challenged the top-down view. Modelling of the reaction time (RT) distributions led them to conclude that increasing the local feature contrast *always* benefited search. According to their model, the target was selected as the first item on 100% of all trials in densely packed displays, irrespective of whether the target colour was kept constant or varied across subsequent trials. They concluded that, in dense displays, high bottom-up feature contrasts immediately led to selection of the target, and that top-down tuning could only modulate target selection in sparse displays. Similarly, with regard to inter-trial priming, Rangelov et al. (2017) found that inter-trial priming could enhance target detection only in sparsely populated displays when the local feature contrast of the target was low. Critically, in sparse displays, target selection rates remained well below 100%: According to their estimates, the target was selected as the first item on 61% of the trials on target colour change trials, and on 80% of trials when its colour was kept constant across trials. Rangelov et al. (2017) concluded that sparse displays produce a qualitatively different search and impose hard bottom-up limitations on selection that cannot be completely overcome by other mechanisms (top-down or priming; see also Sagi & Julesz, 1987).

These conclusions are in line with Meeter and Olivers (2006), who also found feature priming effects in sparse displays (of three items), which were reduced or eliminated in dense displays (with 12 items). Meeter and Olivers (2006) proposed an *ambiguity resolution account* for priming effects, according to which feature priming only modulates performance when the relative saliency of the target is low (i.e., in sparse displays, when the local feature contrast of the target is low) because only this condition has ambiguity concerning the guidance signal.

These results seem to support a bottom-up view, in which local feature contrasts are the most important determiner for target selection. According to the parameter estimates of Rangelov et al. (2017), low local feature contrasts (i.e., sparse displays) limit our ability to select a pop-out target, reducing target selection rates by 20% when the target is constant, and by almost 40% when the target colour varies. Contrary to the findings of Bravo and Nakayama (1992), knowledge of the target (i.e., keeping the target and nontarget colours constant) did not lead to high target selection rates or immunised visual search against influences of local feature contrasts. Rather, low local feature contrasts seemingly limited target selection even when attention was top-down tuned to the target colour (in constant target conditions) or when attention was biased to the target by inter-trial priming (on repeat trials in variable target conditions).

Note that the diverging empirical findings supporting the bottom-up saliency view and the top-down view, respectively, may be due to methodological differences. Bravo and Nakayama (1992) assessed inverse set size effects of RTs, whereas Rangelov et al.'s (2017) results were based on parameter estimates derived from computational modelling.

Critically, however, conclusions based only on RTs and errors (including modelled parameters of RT and error distributions) are problematic, as RTs and error rates do not necessarily reflect early, attention-guiding selection processes, but could reflect later, decisional or response-related effects (e.g., Huang, Holcombe, & Pashler, 2004; Huang & Pashler, 2007; Lamy, Darnell, Levi, & Bublil, 2018; Shiu & Pashler, 1994). To provide a more decisive test of the role of bottom-up feature contrasts in pop-out search, Becker and Ansorge (2013) measured eye movements to distinguish between early versus late processing stages in visual search. Critically, their results showed that increasing the local feature contrast facilitated later, decisional processes (as reflected in target dwell times; i.e., the durations participants fixated on a target), but not early, pre-attentive or attentive processes (i.e., as indexed by the first eye movements to the target or the number of fixations). These results shed doubt on the view that feature contrasts will always affect pre-attentive processes or selection proper and argue against a strong version of the bottom-up saliency view.

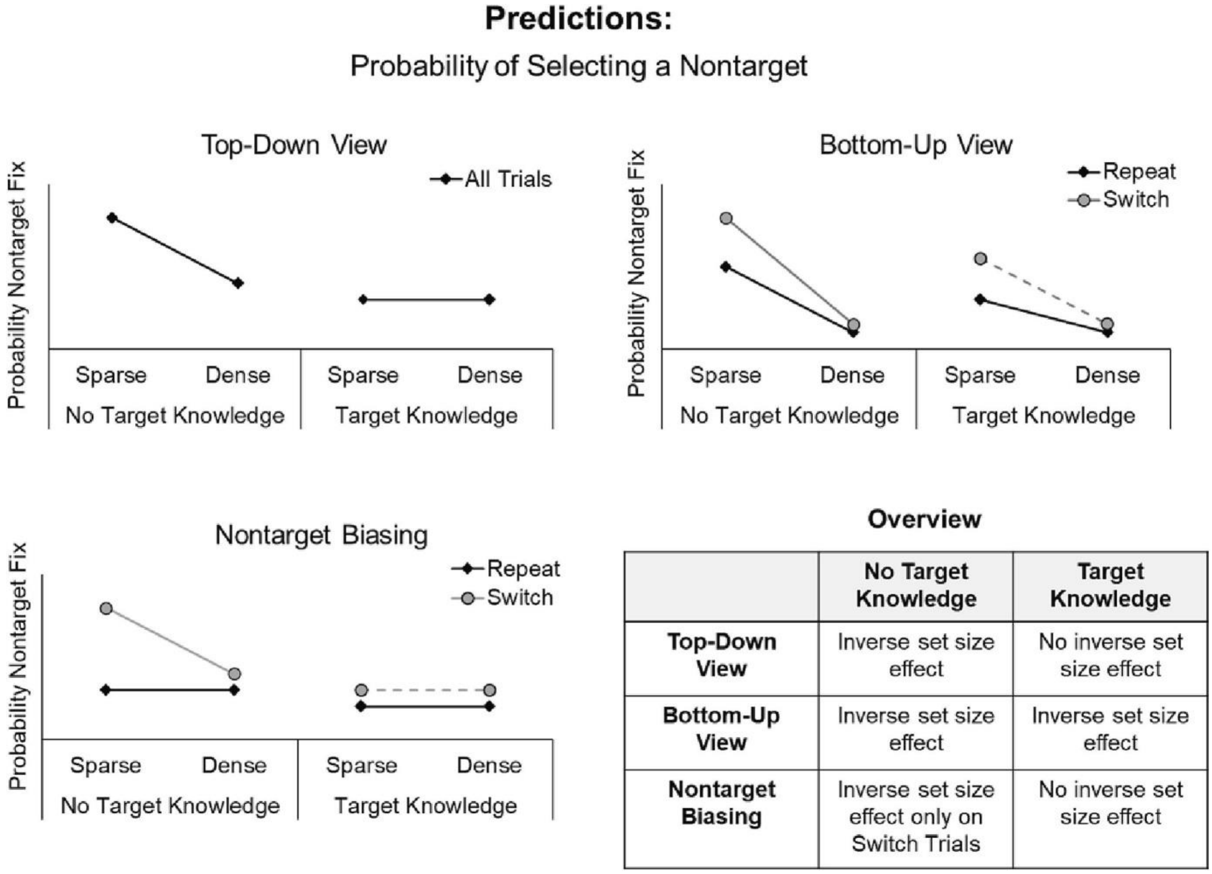
However, Becker and Ansorge's (2013) study is inconclusive, as one major difference between prior studies and this study is that previous experiments directly swapped the target and nontarget colours on switch trials (e.g., Bravo & Nakayama, 1992; McPeek et al., 1999; Meeter & Olivers, 2006; Rangelov et al., 2017), whereas Becker and Ansorge (2013) merely changed the target colour (e.g., between red and yellow), while the nontarget colour was kept constant (e.g., orange). Swapping the target and nontarget features typically leads to stronger priming effects because attention is more strongly biased to the nontarget items after a swap, which in turn leads to higher nontarget selection rates and, consequently, larger switch costs than changing only the target feature (e.g., Kristjánsson & Driver, 2008; Lamy, Antebi, Aviani, & Carmel, 2008; Maljkovic & Nakayama, 1994; but see Becker, Valuch, & Ansorge, 2014).

If the difference in direct swap trials explains the difference in results, it might mean that *local feature contrasts only modulate visual search when attention is biased to the nontargets* (e.g., when the features of the target and nontargets swap). According to this *nontarget biasing* hypothesis, increases in local feature contrasts (increased set size) should aid visual search on target colour change trials, but only when the target and distractor colours were swapped. By contrast, increasing the local feature contrast (i.e., set size) would have no or only a very small effect when attention is already biased to the target (e.g., on repetition trials, or when the target feature is known).

2. Aim of the present study

The aim of the current study was to provide a critical test of the discussed views – the top-down view, bottom-up account, and nontarget biasing hypothesis. To that aim, we measured observer's eye movements in a pop-out search task in which the colours of the target and nontargets (red, green) could randomly repeat or fully switch, compared to the previous trial.

The predictions for each view are depicted in Fig. 1 and are framed in terms of the probability of fixating a nontarget (rather than the target), so that benefits in performance are reflected in lower values and an inverse set size effect (similar to the mean RTs), rather than a positive set size effect (which would result if we plotted the probability of fixating the target).



	No Target Knowledge	**Target Knowledge**
Top-Down View	Inverse set size effect	No inverse set size effect
Bottom-Up View	Inverse set size effect	Inverse set size effect
Nontarget Biasing	Inverse set size effect only on Switch Trials	No inverse set size effect

Fig. 1. Predictions of the top-down view, bottom-up view, and the nontarget biasing account with respect to the probability of selecting a non-salient nontarget when the conditions allow versus do not allow top-down tuning to the target. According to the top-down view, dense displays only benefit search when the defining feature of the singleton target is unknown. According to the bottom-up view, sparse displays should always hamper singleton target selection, regardless of top-down knowledge and feature priming. According to the nontarget biasing account, dense displays only benefit search when attention is biased to the non-salient nontargets – that is, on switch trials, when the defining feature of the singleton target is unknown. The bottom right panel summarises the predictions about inverse set size effects for each account, reflected in the differences between sparse and dense displays. Fix: Fixation.

According to the top-down view (Bravo & Nakayama, 1992), display density should only aid search when the features of the singleton target and non-salient nontargets switch unpredictably and attention cannot be top-down tuned to the target, whereas display density should not affect search when the target is known. The top-down view does not make detailed predictions about priming effects, as priming effects were only discovered later, by Maljkovic and Nakayama (1994). Hence, the prediction graph in Fig. 1 does not distinguish between repeat versus switch trials (see top left panel of Fig. 1).

According to the bottom-up view, increasing the local feature contrast of the singleton target should always benefit search, and sparse displays should always limit the ability to select the target. This should lead to an inverse set size effect across all conditions, independently of target knowledge and priming (i.e., repeat vs. switch trials). Moreover, when the display density is high, target selection should be close to perfect, and selection of the non-salient nontargets should be close to zero, across all conditions (see top right panel of Fig. 1; and Meeter & Olivers, 2006; Rangelov et al., 2017). Priming effects (i.e., advantages for target-repeat relative to target-switch trials) should only occur in sparse displays, when the local feature contrast of the target is low. According to Rangelov et al.'s (2016) parameter estimates, sparse displays should reduce target selection by 20% on repeat trials when the target is known. (The predicted decrement on switch trials is unknown, as reflected in the dashed line for switch trials; see top right panel of Fig. 1).

According to the nontarget biasing account, increasing display density would only benefit search when attention is (strongly) biased to the non-salient nontargets, which would only be the case on switch trials, when (1) the target feature is unknown and (2) the features of target and nontarget swap, compared to the previous trial (see bottom left panel of Fig. 1), because attention would only be fully biased to the nontargets in these conditions. In all other conditions (target knowledge or repeat trials), target selection rates should be high (and nontarget selection should be low) because attention is likely biased to the target due to prior knowledge or priming. The nontarget biasing account would not make any predictions regarding the magnitude of the priming effect (reflected in the dashed line for switch trials with target foreknowledge; see Fig. 1).

As highlighted in the overview table (bottom right of Fig. 1), the main predictions of the three views refer to the presence versus absence of inverse set size effects in the conditions, which reflect benefits in target selection when the display density increases.

3. Overview of experiments

The predictions outlined above were tested in two eye tracking experiments with a colour pop-out search task. In both experiments, we systematically varied bottom-up feature contrasts by varying display density between three, six, and 12 items. In Experiment 1, we varied top-down knowledge by presenting a valid or neutral word pre-cue prior to each trial and varied the trial history by repeating versus switching the target and nontarget colours between red and green.

In Experiment 2, we systematically varied the strength of attentional biasing to the nontargets, by comparing the effects of different kinds of intertrial changes (e.g., target change, nontarget-change, partial swaps) on feature priming effects and inverse set size effects. As in our previous study (Becker & Ansorge, 2013), we used eye tracking to distinguish between early attention-guiding processes and later processes that commence after the target has been found.

4. Experiment 1: Feature contrast, priming and top-down knowledge

Experiment 1 was a classical pop-out search task in which participants had to search for a singleton target with a unique colour (red or green) that was presented among nonsingleton nontargets of the opposite colour (green or red). The search displays contained either three, six, or 12 search items. Participants had to respond to a small arrowhead inside the target to record that they had found the target (see Fig. 2). Moreover, prior to each trial, we presented a *word cue* that was either valid or neutral. Valid pre-cues informed participants of the target colour in the upcoming trial (100% valid), whereas neutral pre-cues consisted of the words “RED OR GREEN” and provided no information about the upcoming target. The target and nontarget colours randomly repeated or switched across trials (Becker, 2008a), and we assessed inverse set size effects and top-down tuning effects separately for repeat and switch trials. The magnitude of the feature priming effects (performance on switch minus repeat trials) was also assessed, separately for each of the set size conditions (3, 6, 12) and pre-cueing conditions (valid, neutral).

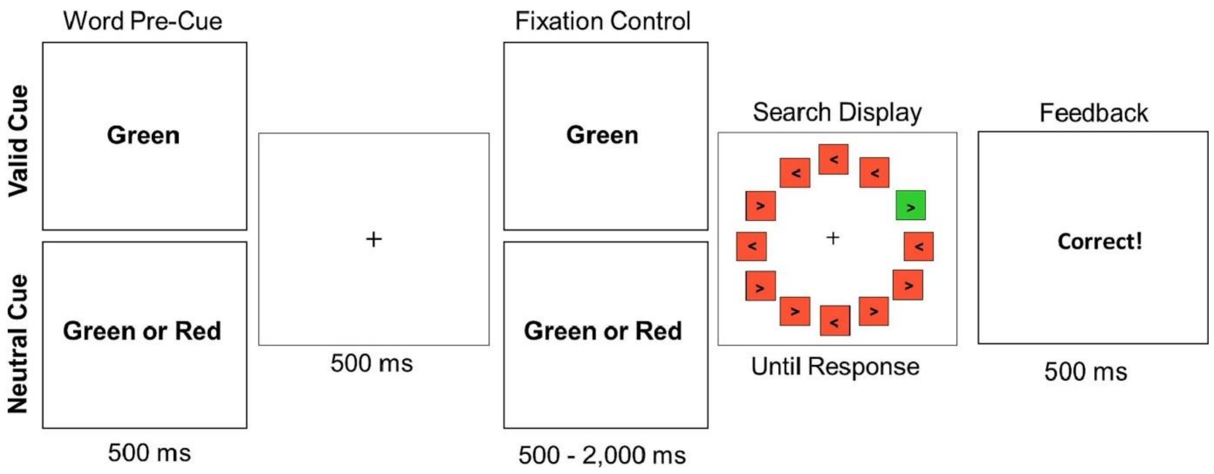


Fig. 2. Schematic example of a trial in Experiment 1. Participants were first presented with a word pre-cue (top: valid cue; bottom: neutral cue), followed by a fixation cross and another presentation of the pre-cue that was yoked to a fixation control. Once participants were fixating on the centre of the word cue, a search display (3, 6, or 12 items) was presented. Participants had to search for the odd-coloured singleton target and respond to the item inside with a button press (>: right; <: left). The response was immediately followed by a feedback display.

As in Becker and Ansorge (2013), we measured eye movements during visual search, to distinguish between processes at an early, intermediate, or late level in visual search. Firstly, to index early attention-guiding processes, we analysed the *proportion of first eye movements* that were mis-guided to any of the nontarget items (cf. Van Zoest et al., 2004). Second, to index intermediate search-related processes leading up to target selection, we recorded the *mean number of fixations* until the target was selected. The mean number of fixations during search most closely reflects how often attention was mis-guided to the non-salient nontargets during a trial and, thus, reflects differences in attentional guidance to the target (cf. Najemnik & Geisler, 2005). Third, to capture processes at a late stage of visual search, we report mean *dwell times* on search items, which is the time the eyes remain fixated on a target or nontarget item. The mean dwell times indicate the time needed to process the selected item to determine whether it is the target or a nontarget, and thus show differences in post-selective processing of the (relevant) feature(s) of the selected item (e.g., Becker, 2010, Becker, 2011; Becker & Ansorge, 2013; see also Horstmann, Becker, & Grubert, 2020).

In addition, we report mean RTs and error scores, which are influenced by both early guidance-related processes as well as later decisional and response selection processes. As such, the mean RTs and errors provide a combined measure of early, intermediate, and late processes in visual search, and allow better comparing of our results to previous studies, which often used mean RTs and errors as the main dependent variables (e.g., Bravo & Nakayama, 1992; Meeter & Olivers, 2006; Rangelov et al., 2017). Readers who would like to focus on one or another type of process or outcome (i.e., in early selection or mean RTs) are invited to skip the analyses of the other dependent variables.

The predictions were as follows: According to the top-down view, feature contrast effects, as indexed by inverse set size effects, should only occur when the target is uncertain (i.e., with a neutral cue), but not when participants know the upcoming target colour in advance (i.e., on valid cue trials; Bravo & Nakayama, 1992). According to the bottom-up view, we would expect feature contrast effects (i.e., inverse set size effects) across all conditions, regardless of whether observers knew the upcoming target colour in advance or not, and regardless of feature priming (i.e., both on repeat and switch trials; e.g., Rangelov et al., 2017). Priming effects, as reflected in better performance on repeat than switch trials, should occur only with sparse displays (i.e., set size 3), but not in dense displays (i.e., set size 12; Meeter & Olivers, 2006; Rangelov et al., 2017). By contrast, according to the nontarget biasing explanation, inverse set size effects should only occur when attention is biased to the nontargets (i.e., on switch trials in the neutral cue condition). Feature contrast should not affect search or produce only very weak effects (i.e., weak inverse set size effects) when attention is biased to the target (i.e., on valid cue trials or when the target is repeated).

According to all accounts, the respective influences highlighted by the accounts – namely feature contrast, top-down tuning, and feature priming – should modulate early attention-guiding processes, so that the predicted effects should already be visible in the proportion of first fixations to the target (or nontargets).

4.1. Method

4.1.1. Participants

We chose the same sample size for the present study as for the related study of Becker and Ansorge (2013; Exp. 1; *n* = 12). Given the reported effect sizes for priming effects (partial $\eta^2 = 0.83$ in Becker & Ansorge, Exp. 1), and top-down tuning (partial $\eta^2 = 0.81$ in Becker et al., 2017, Exp. 2) on the proportion of first eye movements to the target, this sample size (12) should result in a power of 0.95 to detect priming and top-down effect, which are assumedly weaker than the feature contrast effect (G*Power).

Four male and eight female participants (mean age: 21.25 years, range: 18–24), completed Experiment 1 for course credit or monetary compensation ($10). All participants had self-reported normal colour vision and normal or corrected-to-normal visual acuity. The procedures of this and all following experiments were approved by the Ethics Committee of The University of Queensland, Australia.

4.1.2. Apparatus

Stimuli were displayed on a 19" colour monitor with a resolution of 1024 × 784 pixels. The timing of the experiment and response collection were controlled using the software Presentation (Neurobehavioural systems). Responses were collected with a standard USB mouse, and eye movements were monitored with a video-based eye tracker (Eyelink 1000, SR Research, Ontario, Canada) at 500 Hz.

4.1.3. Stimuli

Search displays consisted of three, six, or 12 coloured squares (1.6° × 1.6°) that were equidistantly distributed on the circumference of an imaginary circle with a diameter of 16.9°, against a white background (see Fig. 2 for an example of the search displays). In the set size 3 condition, stimuli appeared either at the 12 o'clock, 4 o'clock and 8 o'clock positions, or at the 2 o'clock, 6 o'clock and 10 o'clock positions. In the set size 6 condition, all of the just-mentioned positions were filled with stimuli, and in the set size 12 condition, the search stimuli were placed at positions from 1 o'clock to 12 o'clock. The squares were coloured red (RGB: 255, 120, 90; Lu'v': 21.7, 0.272, 0.509) or green (RGB: 0, 185, 0; Lu'v': 21.7, 0.120, 0.563), and were adjusted to be equiluminant using a CRS colorimeter. All search items contained a small black < or > symbol (0.2° × 0.2°; Arial Black, 10 pt) as a response-defining item.

4.1.4. Design

The experiment consisted of two blocked conditions. In the *Valid Cue* condition, participants were informed about the colour of the target on the next trial via a word cue (“RED” or “GREEN”; 100% valid; Arial Black, 14 pt). In the *Neutral Cue* condition, the valid pre-cue was replaced with a neutral cue that indicated both possible target colours (i.e., “RED or GREEN”; Arial Black, 14 pt).

Within each block, the target and nontarget colours repeated or switched randomly. The set size condition, target position, and arrow direction were chosen randomly on each trial, with the limitation that each display contained an equal number of leftward and rightward-pointing arrows as response-defining items (except for the set size 3 condition, which could contain two leftward or two rightward-pointing arrowheads). The order of the two blocked conditions was counterbalanced across participants, and each participant completed 600 trials (300 per blocked condition).

4.1.5. Procedure

Testing was completed in a normally lit room, with the participant's head resting on a chinrest and against a forehead-rest of the eye tracker. The monitor-to-head distance was 62 cm. Prior to the experiment, participants were given written instructions about the task and the possible target and nontarget colours in the upcoming block. All observers were instructed to make a fast and accurate eye movement to the odd-coloured target, and to press the right mouse button when the target contained a ">" symbol, and the left mouse button when it contained a "<" symbol.

To ensure stable and accurate eye tracking, participants were calibrated with a nine-point calibration prior to each block and after each break. To ensure that participants read the word pre-cue, the valid and neutral pre-cues were first presented for 500 ms, then disappeared for 500 ms (with only the fixation cross being presented), and then re-appeared for at least 500 ms (for a maximum of 2 s, coupled to a fixation control that only presented the search display when the gaze was within 50 pixels of the centre of the display). The search display was presented until the button press response and immediately followed by a feedback display informing participants about the accuracy of the button press response (using the words “Correct” or “Wrong”), presented for 500 ms. After an intertrial interval of 250 ms, during which a blank white screen was presented, the next trial started, again with the word pre-cue.

All materials and data of Experiments 1 and 2 are publicly available via this link: https://www.dropbox.com/sh/fefvjl8nt5btv7p/AACQJKyCL59yb_lc2PtkO1Zca?dl=0.

4.2. Results

4.2.1. Data

Eye movements were parsed into saccades, fixations, and blinks using the standard parser configuration of the Eyelink software, which classifies an eye movement as a saccade when it exceeds a velocity of 30°/s or an acceleration of $8000°/s^2$. Fixations were assigned to a stimulus (target, nontarget) when the gaze was within 3.2° (100 pixels) of the centre of a stimulus.

We excluded all trials in which participants failed to select the target within 2 s from the onset of the search display, which led to a loss of 3.2% of the data. In addition, we excluded trials with anticipatory responses (RT < 200 ms) or delayed responses (RT > 2 s), which led to an additional loss of 0.2% of the data.

4.2.2. Proportion of first nontarget fixations

The results of the first fixations on each trial are depicted in Fig. 3 (top left panel). We chose to report the proportion of first *nontarget* fixations rather than *target* fixations to render the results more comparable across the different dependent measures (i.e., so that an increase in values always indicates a decrement in performance).

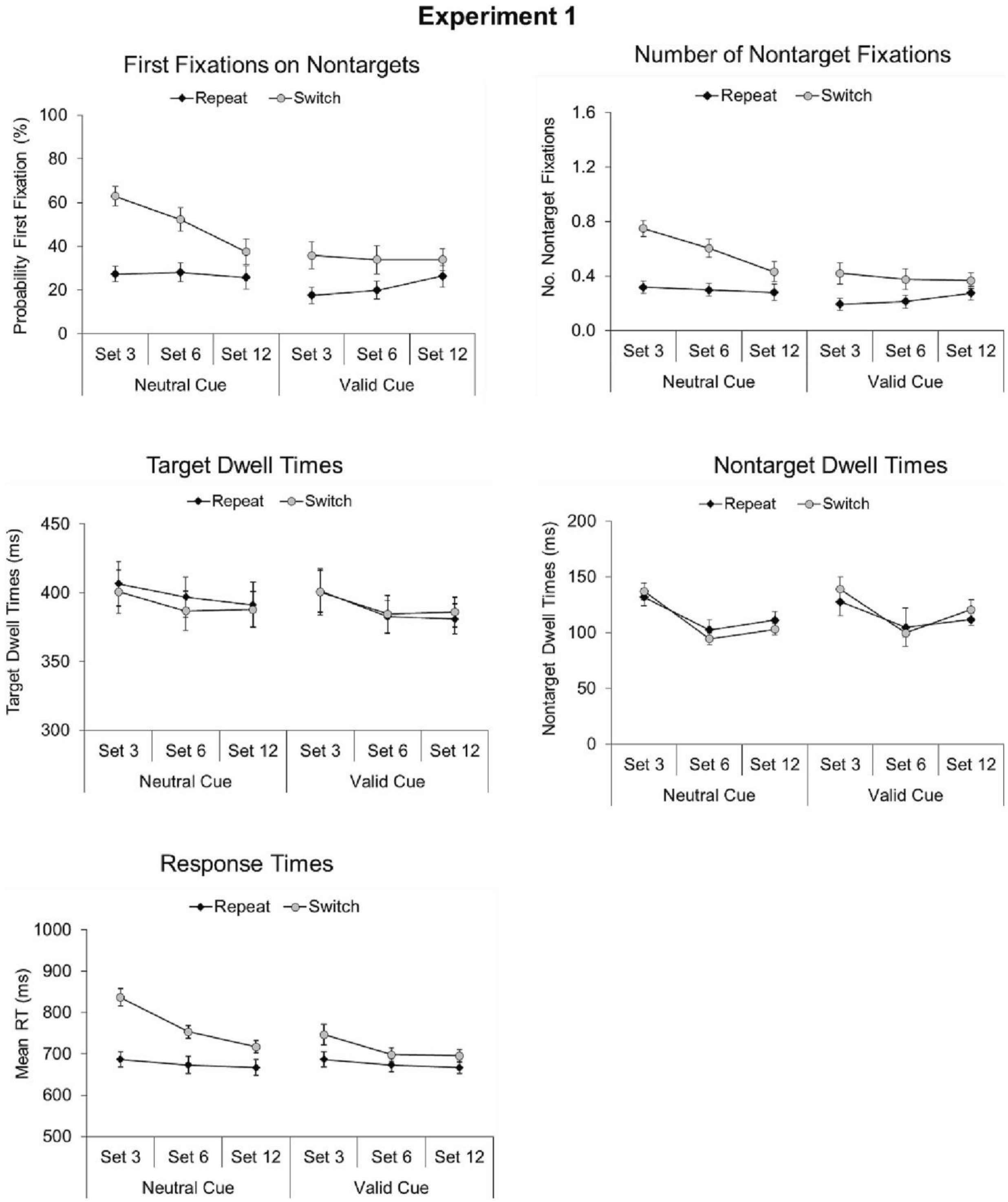


Fig. 3. Results of Experiment 1: The first fixations and number of fixations until target selection both showed strong effects of top-down tuning (cue validity) and priming effects (repeat vs. switch trials). Target feature contrasts modulated search only on switch trials in the neutral cue condition. Later measures, such as the dwell times (middle panels) showed no priming or top-down tuning effects, but reliably showed inverse set size effects. The mean reaction time (RT; bottom left panel) reflected a combination of early and late effects. Apparently, sparse displays do not impair early attention-guiding processes when attention is biased to the target (either via top-down processes or repetition of the target), but hamper later, target identification processes. Error bars depict +/− 1 SEM.

We first computed an omnibus 2 × 2 × 3 analysis of variance (ANOVA) over the first eye movements to the nontargets, with the variables word cue (valid vs. neutral), priming (repeat vs. switch trial) and set size (set size 3 vs. 6 vs. 12). The results showed significant main effects and interactions across all factor combinations (see Table 1), reflecting that bottom-up feature contrast, top-down tuning, and the trial history interacted in the guidance of attention.

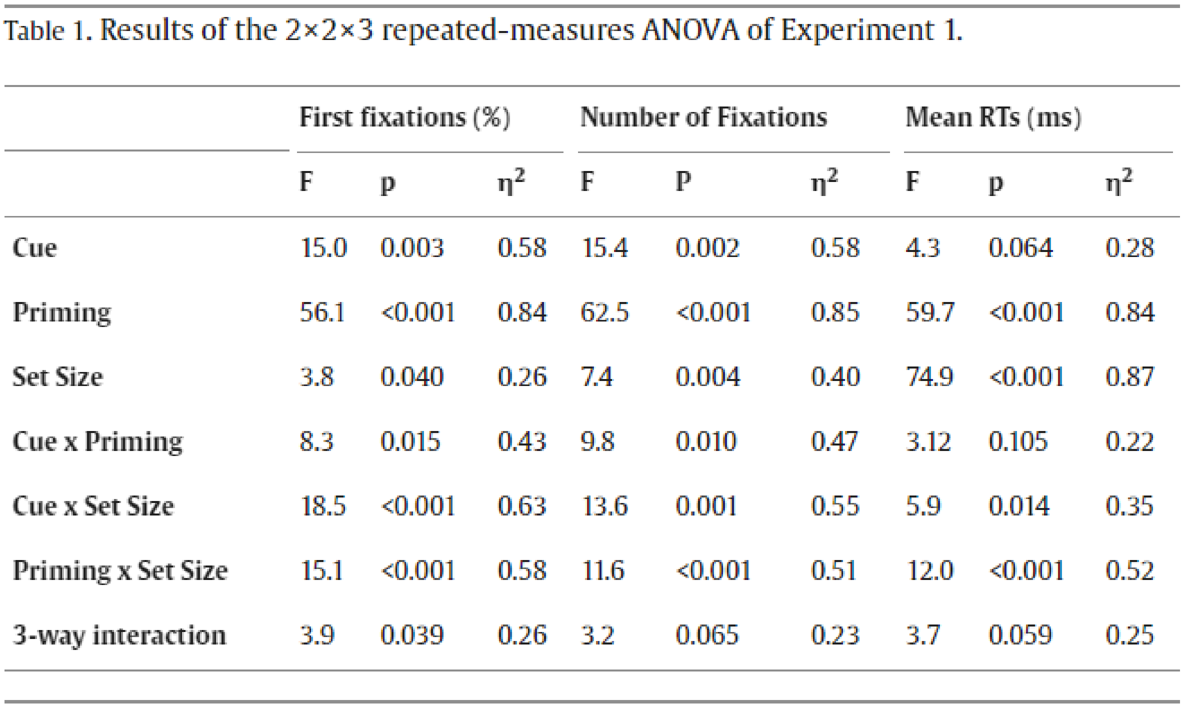


Table 1. Results of the 2×2×3 repeated-measures ANOVA of Experiment 1.

	First fixations (%)			Number of Fixations			Mean RTs (ms)		
	F	p	η^2	F	P	η^2	F	p	η^2
Cue	15.0	0.003	0.58	15.4	0.002	0.58	4.3	0.064	0.28
Priming	56.1	<0.001	0.84	62.5	<0.001	0.85	59.7	<0.001	0.84
Set Size	3.8	0.040	0.26	7.4	0.004	0.40	74.9	<0.001	0.87
Cue x Priming	8.3	0.015	0.43	9.8	0.010	0.47	3.12	0.105	0.22
Cue x Set Size	18.5	<0.001	0.63	13.6	0.001	0.55	5.9	0.014	0.35
Priming x Set Size	15.1	<0.001	0.58	11.6	<0.001	0.51	12.0	<0.001	0.52
3-way interaction	3.9	0.039	0.26	3.2	0.065	0.23	3.7	0.059	0.25

Analysing set size effects in the Neutral Cue condition with one-way ANOVAs separately for the two priming conditions showed a highly significant inverse set size effect on switch trials, $F(2,22) = 24.4$, $p < .001$, $\eta_p^2 = 0.69$, but not on repeat trials, $F < 1.0$ (see Fig. 3, top left).

Critically, the Valid Cue condition revealed no inverse set size effect on switch trials, $F < 1.0$, or on repeat trials, which conversely showed a slightly positive set size effect, $F(2, 22) = 10.4$, $p = .001$, $\eta_p^2 = 0.49$. Thus, in line with both the top-down tuning account and the nontarget biasing account, valid pre-cueing or knowledge of the target colour eliminated the benefits of increasing feature contrast (both on repeat and switch trials). Moreover, in additional support of the nontarget biasing explanation, feature contrasts did not modulate attention when attention was biased to the target, as reflected in the absence of inverse set size effects on repeat trials (both with a valid and a neutral cue).

Second, to test if feature priming effects were eliminated in the higher set size conditions, we also compared priming effects in each set size × pre-cue condition. Priming effects remained significant across all set size and pre-cue conditions, as shown by significant differences between repeat and switch trials in the respective two-tailed *t*-tests; all *t*s > 2.6, *p*s ≤ 0.025 (see Fig. 3, top left panel).

For completeness, we also analysed top-down tuning effects across the conditions. The results revealed that valid pre-cueing significantly reduced selection of the nontargets compared to the neutral cue in the set size 3 and 6 conditions, both on repeat and switch trials, all *t*s > 2.4, *p*s ≤ 0.034, but not in the set size 12 conditions, both *t*s < 1.1, *p*s > 0.32.

4.2.3. Number of nontarget fixations

To assess possible differences in target guidance more comprehensively, we next analysed the number of nontarget fixations prior to target selection. The results of the 2 × 2 × 3 ANOVA showed significant effects and interactions for all variables, with the exception of the three-way interaction which just failed to reach significance (see Table 1).

Analysing the data for inverse set size effects in the Neutral Cue condition showed no effect of set size on repeat trials, $F < 1.0$, but highly significant inverse set size effects on switch trials, $F(2, 22) = 16.4$, $p < .001$, $\eta_p^2 = 0.60$ (see Fig. 3, top right panel). The Valid Cue condition showed a slightly positive set size effect on repeat trials, with increases of the set size leading to performance decrements, $F(2, 22) = 7.6$, $p = .003$, $\eta_p^2 = 0.41$. Validly cued switch trials showed no significant set size effect, $F(2, 22) = 1.5$, $p = .24$, again supporting a nontarget biasing account over a bottom-up view.

The priming effect decreased with increases in the set size, but it remained significant across all set size and pre-cue conditions, all *t*s > 2.8, *p*s ≤ 0.016.

Top-down knowledge in the valid cue condition led to significantly fewer fixations compared to the neutral cue condition in the set size 3 and 6 conditions (on both repeat and switch trials), all *t*s > 2.5, *p*s ≤ 0.027; but not in the set size 12 conditions, both *t*s < 1.4, *p*s > 0.20.

4.2.4. Dwell times

As Becker and Ansorge (2013) found inverse set size effects in a late measure (i.e., the target dwell times), we also assessed dwell times on the target and nontargets. First, a 2 × 2 × 3 ANOVA computed over the *target* dwell times revealed only a main effect of set size, $F(2, 22) = 6.5$, $p = .011$, $\eta_p^2 = 0.37$, but no other main effects or interactions, all Fs < 2.4, ps > 0.14. As shown in Fig. 3 (2nd row, left), the target dwell times showed an inverse set size effect, replicating previous results (Becker & Ansorge, 2013).

The same analysis computed over the *nontarget* dwell times similarly showed only a highly significant inverse set size effect, $F(2,20) = 22.2$, $p < .001$, $\eta_p^2 = 0.69$,[2] but no other main effects or interactions, all Fs < 1.0. These results confirm that a higher number of nontargets in the display facilitates late, target and nontarget identification processes.

4.2.5. Mean reaction times

The mean RTs are depicted in the bottom right panel of Fig. 3. The same 2 × 2 × 3 ANOVA computed over the mean RTs revealed no significant effect of pre-cueing, but significant effects of priming and an inverse set size effect, which was modulated by an interaction with pre-cueing (see Table 1).

Inverse set size effects were found with the valid cue on switch trials, $F(2, 22) = 12.0$, $p = .001$, $\eta_p^2 = 0.52$, but not on repeat trials, $F(2, 22) = 2.5$, $p = .113$, *ns*. With the neutral cue, inverse set size effects were found both on switch trials, $F(2, 22) = 41.7$, $p < .001$, $\eta_p^2 = 0.79$, and on repeat trials, $F(2, 22) = 16.5$, $p < .001$, $\eta_p^2 = 0.60$.

Feature priming effects led to significantly shorter RT on repeat than switch trials across all conditions, all ts > 2.5, ps ≤ 0.028, with the sole exception of the neutral cue set size 12 condition, $t(11) = 1.5$, $p = .15$. Top-down tuning to the target on valid cue trials only led to faster RT in the set size 3 condition, both on repeat and switch trials, ts > 2.7, ps ≤ 0.038, but not in any of the higher set size conditions, all ts < 2.2, ps ≥ 0.058.

4.3. Discussion

Experiment 1 yielded several important results. First, focussing on early attention-guiding, and intermediate search-related processes, there was no evidence for the bottom-up view that low local feature contrasts (i.e., sparse displays) generally limit the ability to select the target (e.g., Rangelov et al., 2017). In fact, search performance was highest (not lowest) in a subset of sparse display conditions (when the target was known and repeated; see Fig. 3, top). In line with the nontarget biasing account and Bravo and Nakayama's (1992) top-down view, an increase in feature contrast did not benefit search when the target was known (i.e., validly pre-cued). However, contrary to the top-down view, uncertainty about the target did not automatically lead to inverse set size effects. Rather, in line with the nontarget biasing account, inverse set size effects were only observed when attention was biased to the nontargets (i.e., on switch trials), not on repeat trials.

Thus, the results provided clear evidence for the nontarget biasing account: Evidently, an increase in local feature contrasts of the target only aids search when (1) the target is unknown and (2) attention is biased to the nontarget feature (i.e., on switch trials). This implies a sequence of processes, whereby attention is first biased to the nontarget colour (e.g., via the target colour on the previous trial). As a consequence, one of the nontargets is selected first, and it is only under these conditions that local feature contrasts can help or hinder visual search performance. Possibly, dense displays can facilitate grouping and rejection of the nontargets ‘as a group’ and, thereby, increase confidence about the target location (e.g., Bacon & Egeth, 1991; Becker, Horstmann, & Remington, 2011; Duncan & Humphreys, 1989). Alternatively, it is possible that we can select multiple objects, especially when they are close together (both when an eye movement is executed; e.g., Venini, Remington, Horstmann, & Becker, 2014; and with covert attention only; e.g., Eimer & Grubert, 2014). Selection of multiple non-targets may be more likely in dense displays because of their greater spatial proximity (e.g., Eriksen & St James, 1986), and would also facilitate non-target rejection processes, thus explaining inverse set size effects.

In conclusion, local feature contrasts do not guide attention in the sense of biasing attention to a salient item, as proposed by the bottom-up view. Rather, increasing the number and density of the items facilitates nontarget rejection (i.e., a decision-making process), thereby, allowing attention to be guided more quickly to the target *after* selection of a nontarget.

With this, the results imply a division of processes, whereby increasing the local feature contrast or set size does not appear to directly bias attention to the target itself. Attention is biased to the target feature either via the trial history (priming) or top-down word cues. In addition, once attention is biased to the target, attention can be guided to the target regardless of the number of nontarget items or variations in feature contrast.

Second, regarding the measures of later processing, the present study mirrors the results of Becker and Ansorge (2013) and shows robust inverse set size effects on target dwell times. Extending on previous findings, the present study also showed inverse set size effects on nontarget dwell times. Interestingly, the dwell times remained unaffected by advance knowledge of the target features and priming. These results provide further evidence of a dissociation between early attention-guiding processes and later target identification (or nontarget rejection) processes, which has also been observed in previous studies (e.g., Hamblin-Frohman & Becker, 2021). It also shows the importance of a separation of measures of early, intermediate, and late processes in visual search that are otherwise indistinguishably lumped together in the composite measure of RT.

As a third important finding, the results of the mean RTs replicated the results pattern reported in previous studies, which originally inspired the bottom-up view: In the neutral uninformative pre-cue condition, inverse set size effects were found across all conditions. Feature priming effects were significant only in the small set size conditions (three and six), but not in the set size 12 condition. Superficially, these results seem consistent with the view that feature contrasts will always modulate search, whereas feature priming can only modulate search in sparse displays, in line with a bottom-up saliency account (e.g., Meeter & Olivers, 2006; Rangelov et al., 2017).

However, the present study clarifies that the robust inverse set size effects in the valid cue conditions and on target repetition trials stem from late target identification processes and not early attention-guiding processes. The mean RTs in the present study were strongly influenced by processes at a later stage in visual search whereas earlier, attentional effects were 'washed out': As processes at a later stage were not susceptible to top-down biases or feature priming, these effects were quite weak in the mean RTs, especially in the higher set size conditions. Thus, the mean RTs did not accurately reflect effects in early attention-guiding processes, but over-estimated feature contrast effects (i.e., inverse set size effects) found in the late measures, while under-estimating top-down tuning and priming effects. Deviating from the mean RTs, early and intermediate processes in visual search (until target selection) revealed stronger effects of top-down tuning and priming than of feature contrast. With this, the results confirm previous theoretical considerations and empirical data that mean RTs are strongly dominated by intermediate and late processes in visual search that are reflected in the number and duration of non-target fixations (Horstmann, Becker, & Ernst, 2017). The fact that the mean RTs are dominated by intermediate and late processes in visual search argues against the standard approach of basing inferences about early attention-guiding processes on mean RTs (e.g., Becker, 2010; Hamblin-Frohman & Becker, 2021; Martin & Becker, 2018).

Another interesting finding of Experiment 1 was that feature priming effects were attenuated but not eliminated in the valid cue conditions, when observers knew the upcoming target feature. The same results have been reported in previous studies on feature priming effects (e.g., Becker, 2008b; Cochrane & Pratt, 2020; Folk & Remington, 2008; Leonard & Egeth, 2008), which showed that feature priming effects are typically not completely eliminated by advance information (except when word cues specified one of two possible targets; e.g., Fecteau, 2007). Even though the degree of top-down tuning was insufficient to eliminate feature priming effects in the present study, it was sufficient to eliminate the inverse set size effect. This indicates that feature contrasts overall have a weaker effect on visual search performance than feature weighting processes arising from top-down tuning or priming.

5. Experiment 2: Different types of inter-trial changes

Experiment 1 provided evidence for a nontarget biasing account, which predicted that inverse set size effects or the ability to benefit from increased local feature contrasts, depend on (1) uncertainty about the target feature, (2) attention being biased to the nontargets, whereby facilitated grouping of the nontargets with higher set sizes then explains faster search (a reduction in switch costs) in higher set size conditions. However, as the evidence for nontarget biasing causing the effects is somewhat indirect, Experiment 2 sought to provide a more direct test of whether the inverse set size effect depends on the degree to which attention is biased to the nontargets.

To that aim, we investigated inverse set size effects across a range of different *types* of switch trials. Previous studies compared switch costs when only the target or nontargets changed (whereas the other item(s) remained the same), and switch costs on *half-switch* trials, in which the nontargets on a trial had the previous target colour and the target a new colour (target half-switch trial), or vice versa, when the target had the same colour as previous nontargets and the nontargets had a new colour (nontarget half-switch trial; e.g., Kristjánsson & Driver, 2008; Lamy et al., 2008). The results showed that switch costs are smallest when only the nontargets change (nontarget change trials), slightly larger when only the target changes (target change trial), and substantially larger on half-switch trials, whereby target half-switch trials produce larger costs than nontarget half-switch trials (e.g., Kristjánsson & Driver, 2008; Lamy et al., 2008; Maljkovic & Nakayama, 1994).

To explain these findings, most models of the intertrial priming effect proposed that the priming effect is due to two combined effects: When the target is selected on a given trial, there is an attentional bias for the target colour, which carries over to the next trial and guides attention to matching colours on subsequent trials. Second, the nontarget colour is inhibited on a given trial, and this nontarget inhibition automatically carries over to the next trial and can bias attention away from this colour (e.g., Kristjánsson & Driver, 2008; Lamy et al., 2008; Maljkovic & Nakayama, 1994). Within these target activation / nontarget inhibition models of the priming effect, target activation is assumed to be slightly stronger than nontarget inhibition, which explains the larger switch costs on target change trials than nontarget change trials, and larger switch costs on target half-switch trials than nontarget half-switch trials (e.g., Kristjánsson & Driver, 2008; Lamy et al., 2008; Maljkovic & Nakayama, 1994; but see Becker, 2010, Becker & Ansorge, 2013; Becker et al., 2014, for a different account).

On target and nontarget change trials, either the target or the nontargets have a new (neutral) colour, so according to the target activation / nontarget inhibition account, there is either only a weak bias against selecting the (repeated) nontargets or a weak bias towards selecting the (repeated) target. As both of these biases work in favour of selecting the target, these trial types should not produce large switch costs. In turn, on half-switch trials, either the nontargets have the colour of the (formerly activated) target or the target has the colour of the (formerly inhibited) nontargets, while the other items have a new (neutral) colour. Both of these types of half-switch trials would therefore result in a selection advantage for the nontargets over the target, and hence, should produce large switch costs.

In line with the target activation / nontarget inhibition accounts, eye tracking studies have shown that switch costs are indeed due to erroneous selection of the nontarget items. On switch trials, participants are far more likely to select one of the nontargets with the first eye movement than when the target and nontarget colours repeat, and this increased probability of selecting the nontarget items predominantly accounts for the switch costs (with only little contribution from later, post-selectional processes; e.g., Becker, 2008a, Becker, 2008b, Becker, 2010, 2013; Becker & Ansorge, 2013; see also Exp. 1 above).

In Experiment 2, we used these insights to test whether inverse set size effects indeed depend on the strength of the attentional bias to the nontargets, by systematically varying the kind of intertrial changes of the target and the nontarget features across trials. The target and nontargets could have one of three possible colours (red, green, blue), which changed such that we could distinguish five different types of trials: Two different kinds of half-switch trials in which the colours of target and/or nontargets partially swapped, two different kinds of change trials, in which only the target *or* nontarget colour changed, respectively, and Repeat trials, in which both the target and nontarget features repeated (see Fig. 4).

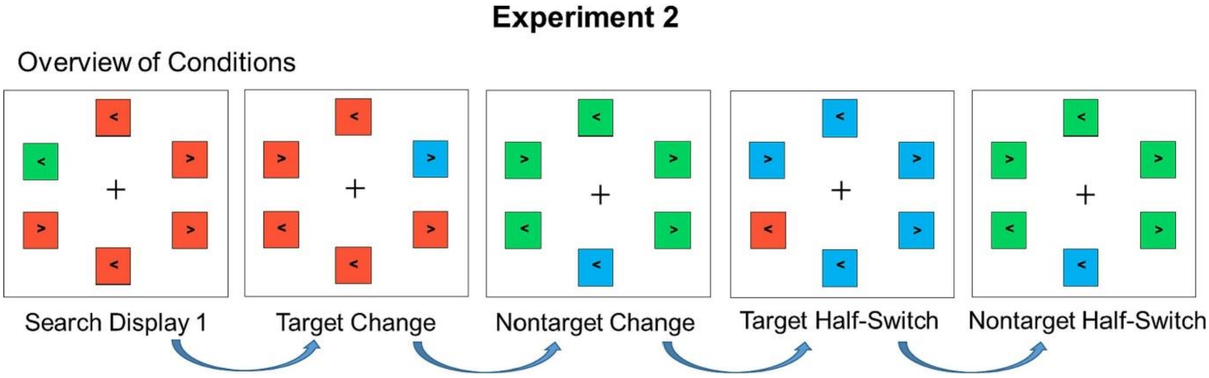


Fig. 4. Overview of the intertrial conditions of Experiment 2. The target and nontarget colours varied randomly between red, green, and blue, creating five different intertrial conditions: Repeat trials, where both the target and nontarget colour repeated; Target and Nontarget Change trials, where only either the target or nontarget colour changed, respectively; and Target and Nontarget Half-Switch trials, in which either the nontargets had the previous target colour (Targ-Hs) or the targets had the previous nontarget colour (Nont-Hs). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Based on previous findings, we expected attention to be most strongly biased to the nontargets on half-switch trials and, especially, when the nontargets inherited the colour formerly associated with the target, while we expected slightly weaker effects when the target inherited the colour previously associated with the nontargets. A much weaker bias to the nontargets should result from changing only the target colour to a new colour, and an even weaker bias when changing only the nontarget colour to a new colour (e.g., Lamy et al., 2008; Maljkovic & Nakayama, 1994).

Importantly, if the magnitude of local feature contrast effects indeed depends on attention being biased to the nontargets, we would expect an inverse set size effect only in the half-switch conditions, and no or a reduced inverse set size effect in the change conditions. Moreover, there should also be no inverse set size effect on repeat trials in the early, attention-guiding measures. A corresponding result would provide more direct evidence for the nontarget biasing hypothesis and establish that feature contrast effects indeed depend on biasing attention to the nontargets.

A corresponding result would also bridge the gap in the previous literature, by explaining why Becker and Ansorge (2013) failed to find inverse set size effects in an early measure when only the target feature changed.

5.1. Methods

5.1.1. Participants

Twelve new participants, four males and eight females (mean age: 21.75 years, range: 19–25), participated in Experiment 2 for course credit or monetary compensation ($10).

5.1.2. Apparatus, stimuli and procedure

These were the same as in Experiment 1, with the following exceptions: First, there were no word cues. Second, all five conditions were presented randomly within a single block that comprised 810 trials. About half of all trials were full repeat trials (both the target and nontarget features repeated), whereas the other half contained approximately equal numbers of trials with different kinds of intertrial changes: On *target change* trials (*Targ-Ch*), only the target had a new colour, compared to the previous trial, whereas the nontarget colour was repeated. On *nontarget change* trials (*Nont-Ch*), only the nontargets had a new colour, whereas the target colour was repeated. Moreover, there were two types of *half-switch* trials: On *target half-switch* trials (*Targ-Hs*), the nontargets had the colour formerly associated with the target, whereas the target had a new colour. On *nontarget half-switch* trials (*Nont-Hs*), the target had the colour formerly associated with the nontargets, while the nontargets had a new colour (see Fig. 4).

5.2. Results

5.2.1. Data

Excluding trials where participants had failed to select the target led to a loss of 2.3% of the data and excluding trials with anticipatory responses (< 200 ms) or delayed responses (> 2 s) led to a further loss of 0.1% of the data.

5.2.2. Proportion of first nontarget fixations

The results of the first fixations in Experiment 2 are depicted in Fig. 5 (top left panel). First, a 5 × 3 ANOVA comprising the variables intertrial condition (Repeat, Targ-Ch, Nont-Ch, Targ-Hs, Nont-Hs) and set size (3, 6, 12) revealed significant main effects of condition, $F(4, 44) = 36.5$, $p < .001$, $\eta_p^2 = 0.77$, and set size, $F(2, 22) = 16.5$, $p < .001$, $\eta_p^2 = 0.60$, as well as a significant Condition × Set Size interaction, $F(2, 22) = 7.3$, $p < .001$, $\eta_p^2 = 0.40$.

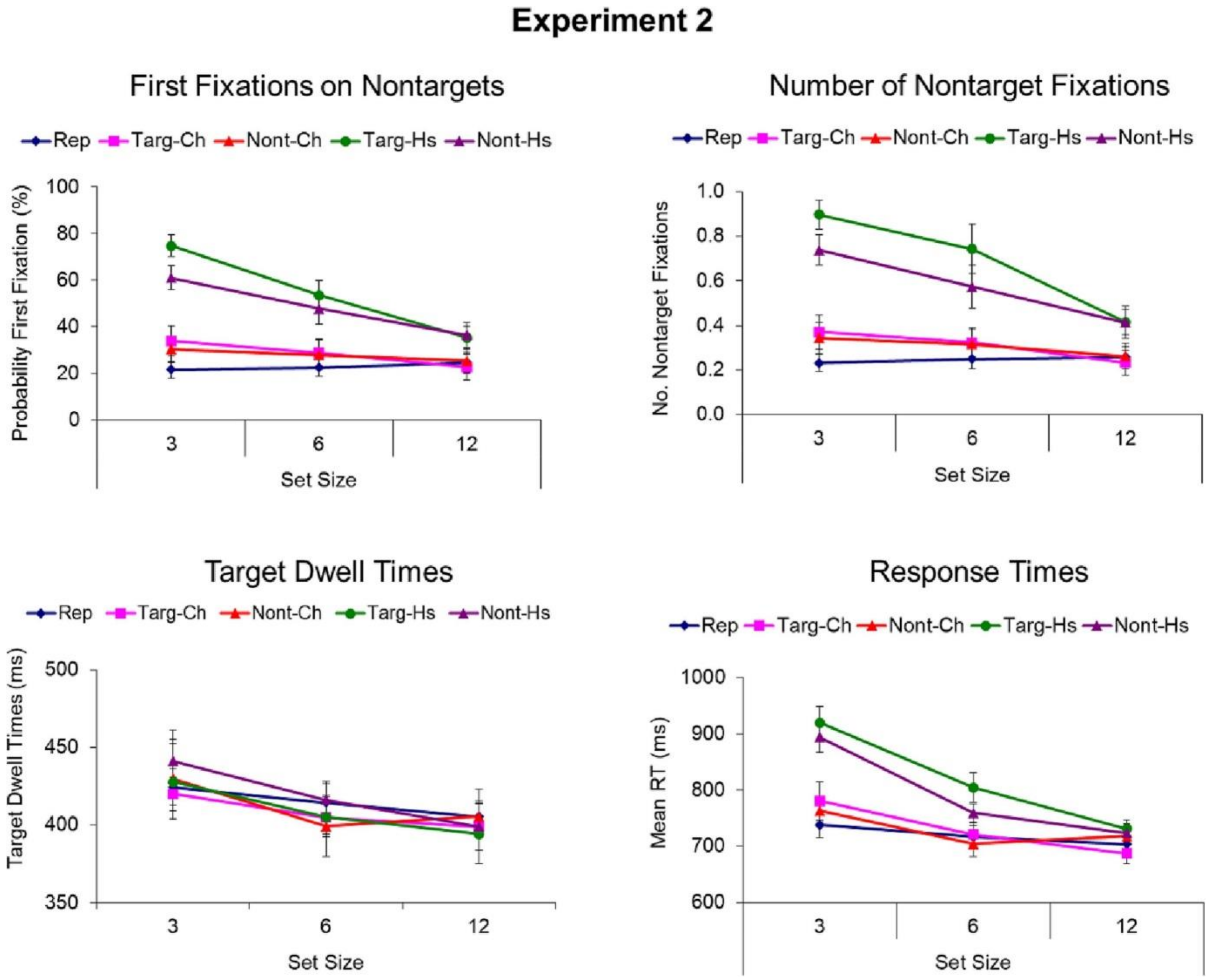


Fig. 5. Results of Experiment 2, depicted separately First Fixations (top left) indexing early processes in visual search, Number of Nontarget Fixations (top right), indexing intermediate processes of target guidance, Target Dwell Times (bottom left), indexing late processes of target identification, and the Response Times (bottom right), which combine all measures. Early and intermediate measures (top graphs) show clear evidence for feature contrast effects and priming effects on half-switch trials, whereas repeat trials and change trials do not show feature contrast effects or consistent priming effects. Late measures (bottom left)

yielded inverse set size effects across all measures, but no or only weak priming effects. The mean RTs (bottom right) show the combined effects, whereby the inverse set size is exaggerated and priming effects are attenuated, due to the stronger influence of later processes. Error bars depict +/− 1 SEM. RT = Reaction Time; Targ = Target; Nont = Nontarget; Ch = Change; Hs = Half-Switch.

Regarding possible set size effects, a series of one-way ANOVAs showed no significant inverse set size effect for repeat trials, $F < 1.0$, target change trials, $F(2, 22) = 1.7$, $p = .21$, or nontarget change trials, $F < 1.0$. Conversely, highly significant inverse set size effects were found on half-switch trials, both when the nontargets inherited the previous target colour (Targ-Hs), $F(2, 22) = 26.1$, $p < .001$, $\eta_p^2 = 0.70$, and when the target had the feature previously associated with the nontargets (Nont-Hs), $F(2, 22) = 10.3$, $p = .001$, $\eta_p^2 = 0.48$.

Priming effects were assessed by comparing the proportion of first fixations in each of the change trials with the repeat trials in the corresponding set size conditions. The results of the pair-wise comparisons showed no priming effects for target change trials, across all set size conditions, all ts ≤ 2.1, ps ≥ 0.56. On nontarget change trials, priming was significant only in the set size 3 condition, $t(11) = 3.0$, $p = .012$, not in the higher set size conditions, all other ts ≤ 1.5, ps ≥ 0.15. In turn, priming effects were highly significant across all set size conditions on both target and nontarget half-switch trials (Targ-Hs and Nont-Hs), all ts ≥ 4.3, ps ≤ 0.001. These results support the view that inverse set size effects only occur when attention is biased to the nontarget feature, which requires swapping the target colour so that it becomes the nontarget colour or vice versa.

5.2.3. Number of nontarget fixations

The number of nontarget fixations in Experiment 2 closely followed the results pattern of the mean proportion of first fixations (see Fig. 5, top right panel). The same 5 × 2 ANOVA showed significant differences between the five intertrial conditions, $F(4, 44) = 43.8$, $p < .001$, $\eta_p^2 = 0.80$, and set size conditions, $F(2, 22) = 19.1$, $p < .001$, $\eta_p^2 = 0.63$, as well as a significant interaction, $F(2, 22) = 6.5$, $p < .001$, $\eta_p^2 = 0.37$.

There were no significant inverse set size effects on repeat trials, $F < 1.0$, target change trials, $F(2, 22) = 2.4$, $p = .13$, or nontarget change trials, $F(2, 22) = 1.9$, $p = .17$. Conversely, highly significant inverse set size effects were found on half-switch trials, both when the nontargets inherited the previous target feature (Targ-Hs), *F*(2, 22) = 10.4, $p = .001$, η_p^2 = 0.49, and when the target had the feature previously associated with the nontargets (Nont-Hs), *F*(2, 22) = 17.1, $p < .001$, η_p^2 = 0.61.

There were no priming effects on target change trials, compared to repeat trials, in any of the set size conditions, all *t*s < 2.1, *p*s > 0.06. On nontarget change trials, priming was significant only in the set size three condition, which differed significantly from repeat trials, *t*(11) = 2.7, $p = .021$ (other *t*s < 1.9, *p*s > 0.08). In turn, priming effects were significant across all set size conditions on both types of half-switch trials (Targ-Hs and Nont-Hs), which both differed significantly from performance on repeat trials, all *t*s > 3.9, *p*s ≤ 0.002.

5.2.4. Dwell times

First, analysing the *target* dwell times with the same 5 × 3 ANOVA revealed only a significant inverse set size effect, *F*(2, 22) = 16.3, $p < .001$, η_p^2 = 0.60, all other *F*s < 1.8, *p*s > 0.16, mimicking the results of Experiment 1. The *nontarget* dwell times showed the same trends as the target dwell times. However, as participants made only a few fixations on the nontargets (≤ 5 fixations per cell for $n = 8$), we did not report the nontarget dwell times.

5.2.5. Mean reaction times

Analysing the mean RTs with the same 5 × 3 ANOVA showed significant main effects of intertrial condition, *F*(4, 44) = 38.4, $p < .001$, η_p^2 = 0.78, set size, *F*(2, 22) = 81.4, $p < .001$, η_p^2 = 0.88, and a significant interaction, *F*(8, 88) = 11.7, $p < .001$, η_p^2 = 0.52.

One-way ANOVAs computed over the mean RTs of each intertrial condition showed significant inverse set size effects across all conditions; on repeat trials, *F*(2, 22) = 16.4, $p < .001$, η_p^2 = 0.60, target change trials, *F*(2, 22) = 11.0, $p = .002$, η_p^2 = 0.50, nontarget change trials, *F*(2, 22) = 9.3, $p = .003$, η_p^2 = 0.45, target half-switch trials, *F*(2, 22) = 47.0, $p < .001$, η_p^2 = 0.81, and nontarget half-switch trials, *F*(2, 22) = 61.5, $p < .001$, η_p^2 = 0.85 (see Fig. 5, bottom).

Priming effects (assessed by comparing the mean RTs in each of the change conditions to performance on repeat trials) were not significant on target change trials or nontarget change trials, across any of the set size conditions, all *t*s < 2.2, *p*s > 0.05. In turn, priming effects were significant across all set size conditions on target half-switch trials, all *t*s > 2.9, *p*s ≤ 0.015.

On nontarget half-switch trials, priming effects were evident in the set size 3 condition, *t*(11) = 9.9, *p* < .001, and the set size 6 condition, *t*(11) = 3.7, *p* = .004, whereas they just failed to reach significance in the set size 12 condition, *t*(11) = 2.2, *p* = .055.

5.3. Discussion

The results of Experiment 2 support the nontarget biasing hypothesis that local feature contrasts only influence attentional guidance when attention is biased to the nontargets. Again, there was no evidence that sparse displays present a bottom-up limitation for efficiently selecting the target, contrary to the bottom-up view. Local feature contrasts had no effect on target selection on repeat trials when attention was biased to the target. On target change and nontarget change trials, increasing the number of nontargets showed a trend for facilitating search, but did not result in a significant inverse set size effect. Significant inverse set size effects only emerged in the half-switch conditions, where attention was biased to the nontargets.

As in previous studies, attention was only biased strongly towards the nontargets when the target and nontarget features partially swapped (as reflected in the significant priming effects compared to repeat trials in these conditions). The fact that inverse set size effects were only observed in those conditions that also showed significant priming effects indicates a direct relationship between the two variables: Feature contrasts only aid search when attention is biased to the nontargets.

These results argue against the common notion that local feature contrasts actively bias attention to the target and instead support the nontarget biasing account, that dense displays facilitate nontarget rejection. In other words, the results suggest a sequence of processes in which attention is first biased to the nontargets (e.g., due to priming), which, secondly, instigates processing of the nontargets, which is facilitated by dense displays (possibly due to a grouping mechanism), and finally, allows localising the target earlier.

Interestingly, later processes as indexed by target dwell times again showed a different results pattern: We found inverse set size effects even on repeat trials when the target was found immediately, and the inverse set size effect did not vary across conditions or as a function of how quickly the target could be found. These results suggest that target identification processes depend on the number of nontargets present in the search display. Target identification is enhanced when the target is dissimilar from a large number of nontargets, and distractor rejection and identification is enhanced when the distractor is similar to a large number of distractors (e.g., Becker, 2011; see also Buetti, Cronin, Madison, Wang, & Lleras, 2016). One possible reason for this facilitation is that, in pop-out or singleton search, there is more evidence that the target is in fact the target when it is dissimilar from a larger number of items; and there is more evidence that a nontarget is in fact a nontarget when it is similar to a larger number of non-salient nontargets (e.g., Becker, 2011; see also Buetti et al., 2016). A higher degree of certainty (or confidence) that the selected item is in fact a target (or nontarget) would translate to shorter dwell or decision times, whereas a lower degree of certainty or confidence would translate to longer dwell or decision times (e.g., because of additional verification steps).

The mean RTs were again influenced by both earlier and later processes: Due to effects in later target identification processes, we found significant inverse set size effects across all intertrial conditions, including repeat trials. The mean RTs also showed modulation by earlier attention-guiding processes in that inverse set size effects were more pronounced in the half-switch conditions than the repeat and change conditions.

In this respect, the mean RTs also support the main conclusion, that inverse set size effects depend on feature priming, and more specifically, of attention being initially guided to the nontargets. However, the significant inverse set size effect across all intertrial types, including on repeat trials, would still wrongly suggest that early selection in visual search always profits from increases in local feature contrast. Hence, the results of Experiment 2 also highlight the need to use measures that tap into early, attention-guiding processes in assessing the true origin of possible feature contrast effects.

6. General discussion

The present study examined the effects of top-down tuning, feature priming, and feature contrast in visual search for a pop-out (or singleton) target, to critically test three different hypotheses about the interplay of top-down and bottom-up processes, and priming effects.

Contrary to the bottom-up saliency view, we found that local feature contrasts did not always affect attention, but only sped up search when attention was biased to the nontargets. Both informing observers about the target colour in the next trial and simply repeating the target and nontargets was sufficient to allow efficient target selection, even in sparse displays where the target's local feature contrast is low.

Proponents of the bottom-up saliency view previously argued that sparse displays render a pop-out target ambiguous (Olivers & Meeter, 2006) and estimated that selection rates for a pop-out target are reduced by 20–40% in sparse displays compared to dense displays (e.g., Rangelov et al., 2017). While these estimates were based on mean RTs and errors, the present study measured attentional guidance more directly, by tracking participants' eye movements (e.g., Becker, 2010; Deubel & Schneider, 1996). We found no evidence for the claim that initial target selection rates generally increase in denser displays. In Experiment 1, comparing only repeat trials in the set size 3 (sparse display) condition with the set size 12 (dense display) condition showed an 8.7% *decrease* in target selection rates in the dense displays for the valid cue condition, and a 1.4% increase in target selection rates in the neutral cue condition. In Experiment 2, repeat trials again showed a 3.1% *decrease* in target selection rates for dense displays, contrary to the bottom-up saliency view. Results consistent with the estimates of Rangelov et al. (2017) were only found in the full switch condition of Experiment 1 with a neutral cue (dense displays led to a 25.4% increase in target selection rates compared to sparse displays) and in the half-switch conditions of Experiment 2 (target selection rates were 39.4% and 24.6% higher in dense than sparse displays, for the target and nontarget half-switch conditions, respectively). These results clearly show that feature contrasts do not always modulate attention. Contrary to the bottom-up saliency view (e.g., Theeuwes, 2013), sparse displays do not act as a bottom-up limitation on target selection. Rather, biasing attention to the target feature either via top-down tuning or priming is sufficient to eliminate any effects of local feature contrast.

As display density completely failed to modulate attention on repeat trials and pre-cued trials, we also cannot claim that local feature contrasts drive any other effects (e.g., priming effects), as proposed in the ambiguity resolution account (e.g., Meeter & Olivers, 2006).

Some may be tempted to argue that the results are still consistent with a bottom-up view, as feature priming effects are in fact reduced as display density increased. It could be argued that high bottom-up feature contrasts (in set size 12) immediately lead to a ceiling effect in performance, which prevents top-down factors and priming effects, so that bottom-up feature contrast dominates feature priming and top-down factors in *this* sense. However, this interpretation is only consistent with the RT results, which indeed showed the highest performance in the set size 12 condition.[3] The bottom-up view is not consistent with the results of the early attention-guiding measures, such as the proportion of first fixations: First, on two occasions, the early attention-guiding measures showed a *positive* set size effect on repeat trials, one of which was significant. This is difficult to reconcile with the view that there is an inverse set size effect underlying target selection across all conditions, as claimed by the bottom-up view. Second and related, performance was not at ceiling in the high contrast, set size 12 conditions: Performance was significantly better in the sparse, set size 3 condition with a valid pre-cue on repeat trials. Third, priming effects were not eliminated, but still present in the high contrast, set size 12 conditions (see Fig. 3). Fourth, in Experiment 2, we directly manipulated the degree to which attention was biased to the nontargets and found that inverse set size effects scaled directly with intertrial changes that biased attention progressively more strongly to the nontargets. It would be difficult to explain this association between intertrial changes and inverse set size effects without acknowledging that the inverse set size effect is driven by priming. In short, the results of the early, attention-guiding measures are inconsistent with bottom-up views. Instead, the data indicate that inverse set size effects occur only when attention is biased to the nontargets, whereby the inverse set size effect itself is probably driven by facilitated nontarget grouping.

Proponents of the bottom-up view may also be tempted to argue that the target was never salient enough in the present displays with a maximum set size of 12 items and that the set size would need to be increased to 36 items (e.g., as in Rangelov et al., 2017) to find the ultimate maximum target selection rates. This line of reasoning is also not convincing: As noted above, the early attention-guiding measures often showed a trend towards a positive set size effect on repetition trials (see Fig. 3, Fig. 5), rather than an inverse or null set size effect, rendering it rather unlikely that further increasing the set size would suddenly result in a reversal of the (positive) set size function. Second, previous studies supporting the bottom-up view also only varied the set size up to 12 items and found inverse set size effects (e.g., Meeter & Olivers, 2006). Third and most importantly, the critique overlooks that the present study replicates the results that originally inspired the bottom-up saliency view. In line with previous studies, we found reliable inverse set size effects in the mean RTs across all conditions.

This persistent inverse set size effect in the mean RTs and the elimination of priming effects in dense displays also originally inspired the bottom-up view, or was cited in support of the bottom-up view (e.g., Meeter & Olivers, 2006; Rangelov et al., 2017). However, we were able to demonstrate that this particular result pattern was not due to early attention-guiding processes, but to later target and nontarget identification processes, which reliably show inverse set size effects and no priming effects (see also Becker & Ansorge, 2013). As these processes commence after the target has been selected (e.g., reflected in the target dwell times), the RT effects cannot be cited in favour of a particular attention-guiding mechanism. Rather than showing ambiguity in early, attention-guiding processes (e.g., Meeter & Olivers, 2006), the results suggest a possible ambiguity in later processes of identifying targets and nontargets (see below for details).

In turn, the results provided partial support for Bravo and Nakayama's (1992) original interpretation of the results, that an increase in feature contrast only benefits search when the target is uncertain (see also Leonard & Egeth, 2008). In line with their top-down view, we found that advance information about the target colour reliably eliminated inverse set size effects and produced flat or slightly positive set size effects. Their top-down explanation is still only partially correct, however, because we identified that it is not target uncertainty that drives feature contrast effects, but an attentional bias to select the nontargets. If target uncertainty would limit the ability to detect the target in low contrast displays and bring about inverse set size effects, we should have observed inverse set size effects with the neutral cues in Experiment 1 and across all conditions of Experiment 2. As the target could have three different colours in Experiment 2, we would have even expected inverse set size effects to be stronger in this experiment than Experiment 1, where the target could only have two different colours. Contrary to these predictions, inverse set size effects were limited to those conditions and trials where attention was biased to the nontargets. Thus, while Bravo and Nakayama (1992) were correct in claiming that top-down knowledge can eliminate inverse set size effects, they did not correctly identify the role of attentional biases in driving inverse set size effects (which could be due to feature priming effects being unknown at that time; first reported by Maljkovic & Nakayama in 1994).

It may be surprising that dynamic feature weighting processes involved in top-down tuning and feature priming can modulate feature contrast effects, which have been proposed to be rooted in hard-wired connections (e.g., inhibitory lateral connections or isofeature suppression, e.g., Itti et al., 1998; Rangelov et al., 2017). However, it should be noted that we propose an alternative explanation for inverse set size effects, which is based on nontarget grouping or – more generally – facilitated nontarget rejection (e.g., by the ability to select multiple items in a limited area of space) rather than processes that yield a clear saliency signal. Whether grouping / nontarget rejection processes are due to hard-wired processes is an open question. However, grouping itself is presumably an automatic process that does not require top-down knowledge (e.g., Duncan & Humphreys, 1989; Moore & Egeth, 1997).

Moreover, even if we assume that inverse set size effects are mediated by hard-wired connections, it would still be plausible that feature contrast effects may typically – that is, without a bias towards the nontarget features – occur slightly later because mutual inhibition requires that neurons fire *in response to the stimulus input*, whereas top-down modulation and the trial history bias attention and eye movements *prior* to the presentation of the search stimuli (e.g., Becker et al., 2017; Exp. 2), possibly by lowering the threshold for firing in corresponding cells (e.g., Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999; see also Bichot, Rossi, & Desimone, 2005; Conway, Hubel, & Livingstone, 2002; Martinez-Trujillo & Treue, 2004).

Fig. 6 provides an illustration of the processes leading to inverse set size effects: If a red target was successfully selected on Trial 1, attention had to be biased towards red (and/or against the non-target colour green) to select the target. These attentional gain settings automatically carry over to the next trial (or simply remain in place; e.g., Becker & Horstmann, 2009; Mueller, Heller, & Ziegler, 1995) and bias attention towards red and (to a lesser extent) against green. When the target on the next trial is green, attention is now biased towards the (red) *nontargets* and *against* the (green) *target*, which will result in selection of a nontarget rather than the target. Identifying the selected item(s) as a nontarget is faster when more nontargets are present in the display (or when more nontargets are in the vicinity), which leads to faster nontarget rejection and earlier target localisation in dense displays than in sparse displays (e.g., because of facilitated nontarget grouping or processing of multiple nontarget items; Duncan & Humphreys, 1989; Venini et al., 2014).

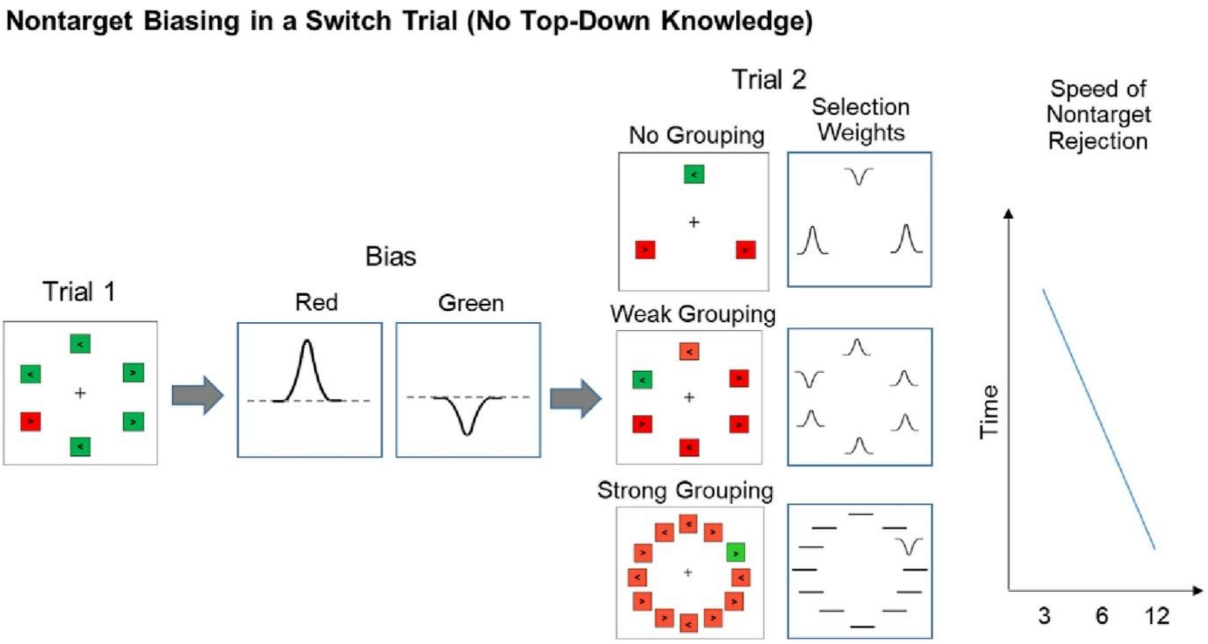


Fig. 6. Illustration of the core ideas of the nontarget biasing account: Selecting the target in Trial 1 leads to an attentional bias towards red and inhibition of the nontarget colour (green). If the target is green on Trial 2, these biases translate into a negative selection weight for the target and a positive bias towards the nontargets. This positive bias towards the notargets decreases with an increase in set size, leading to faster nontarget rejection when the set size is high and an inverse set size effect (right panel). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The eye tracking data directly show facilitation of nontarget rejection in the dwell times, both in this and previous studies (Becker & Ansorge, 2013). It is likely that covert attentional selection will show the same facilitation effect for dense displays, which can cancel a pre-planned eye movement and prevent selection of the nontarget (if covert target localisation is swift enough), thus, explaining how facilitated nontarget rejection can affect the first eye movement on a trial (see Fig. 3, Fig. 5; for saccadic programming and cancelling of saccades, see e.g., McPeek & Keller, 2002; McPeek, Skavenski, & Nakayama, 2000; Theeuwes, Kramer, Hahn, & Irwin, 1998; Walker & McSorley, 2006). Density of similar nontargets might, thus, also decrease search times by decreasing the number of saccades and fixations (cf. Hulleman & Olivers, 2017).

While clarifying the underlying mechanism for feature contrast effects would require further research, display density clearly only plays a substantial role when attention is initially mis-guided to select the nontargets. Similar conclusions were drawn in previous studies. For instance, Bacon and Egeth (1991) disentangled proximity effects from set size effects in pop-out search, and found that inverse set size effects were driven by nontarget grouping rather than target-nontarget proximity. However, in line with previous studies on nontarget grouping (e.g., Duncan & Humphreys, 1989), grouping was supposed to occur at a pre-attentive stage of processing and guide visual attention (Bacon & Egeth, 1991). Similarly, Buetti et al. (2016) proposed that an inverse set size effect does not result from isofeature suppression but rather from a decision-making process. According to Buetti et al.'s (2016) account, the visual system always processes all stimuli in the display (parallel stage-processing) to determine which colour is repeated in more than one item to extract the target location (which gets passed on to the second stage of processing).

However, the present data only partially support these conclusions: Evidence for the claim that the nontargets are always processed was only found in a relatively late measure – the dwell times and mean RTs, not in early and intermediate measures that reflect attention-guiding processes. Early measures only showed evidence for nontarget processing when attention was initially biased to the nontarget feature. As such, Buetti's ‘nontarget processing’ account and previous nontarget grouping explanations do not seem to be applicable to the present results or to pop-out searches that allow biasing attention to the target (either via top-down tuning or history effects).

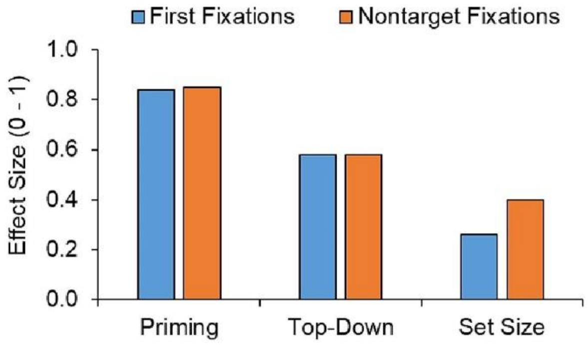
7. Interplay and dominance

One overarching aim of the present study was to shed light on the interplay between top-down tuning, bottom-up feature contrast, and trial history effects, and their possible interactions. We can try to gauge the relative importance or contributions of each of these attentional systems by assessing either (1) which effect is most consistently present (across all conditions and experiments), or (2) by comparing the effect sizes of the main effects.

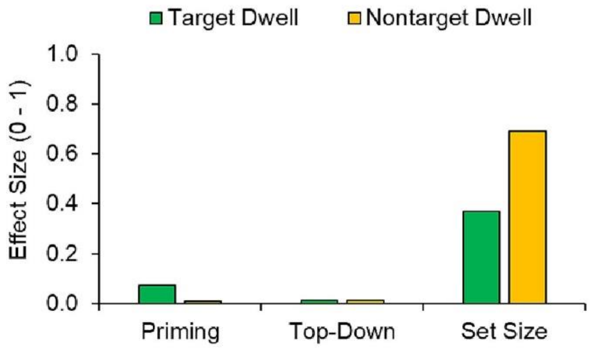
Fig. 7 provides an overview of the prevalence (or consistency) and the effect sizes of bottom-up feature contrast, top-down tuning, and trial history effects in Experiment 1. Both measures (prevalence and effect sizes) suggest that the trial history was the most important determiner of visual selection, followed by top-down tuning, and last, local feature contrasts (see Fig. 7, top graphs).

Effect Sizes

Attention Guiding



Post-Selection



Prevalence

Attention Guiding

	First Fixations	Nontarget Fixations
Priming	6 out of 6 conditions = 100%	6 out of 6 conditions = 100%
Top-Down	4 out of 6 conditions = 66.6%	4 out of 6 conditions = 66.6%
Set Size	1 out of 4 conditions = 25%	1 out of 4 conditions = 25%

Post-Selection

	Target Dwell	Nontarget Dwell
Priming	0 out of 4 conditions = 0%	0 out of 4 conditions = 0%
Top-Down	0 out of 4 conditions = 0%	1 out of 4 conditions = 0%
Set Size	0 out of 4 conditions = 0%	3 out of 4 conditions = 75%

Fig. 7. The effect sizes (left) and prevalence (right) of priming effects, top-down knowledge, and bottom-up saliency (set size) in early measures of visual selection (top) and later, post-selection measures (bottom) in Experiment 1. The depicted effect sizes are the partial eta-squared values of the main effects of the corresponding ANOVAs. Prevalence is computed as the number of conditions in which a given effect reached statistical significance (as per two-tailed *t*-test or one-way ANOVA, within each condition, $p < .05$). The results show that early selection is most strongly affected by priming effects, followed by top-down knowledge, and bottom-up saliency, while later post-selection measures are most strongly affected by bottom-up saliency (as reflected by inverse set size effects).

As the effect size can depend on the strength of a manipulation, and prevalence may be influenced by the particular design and procedures, we should ask if our methods may have led to an under-estimation of some effects. For instance, we found that top-down tuning to the target did not completely eliminate feature priming effects but only reduced them. This finding is in line with multiple previous studies that found that priming effects were reduced but not eliminated when the features switched in a completely predictable sequence or when advance information was provided by 100% valid word cues (e.g., Becker, 2008b; Gaspelin, Gaspar, & Luck, 2019; Hillstrom, 2000; Maljkovic & Nakayama, 1994).

However, Fecteau (2007) has argued that these studies do not provide a fair test of whether top-down tuning can override priming effects because the task can be successfully completed without processing any of the word cues (because the target is un-ambiguously defined as the *odd man out*). Fecteau (2007) also showed that feature priming effects are completely eliminated when each display contains two possible candidate targets, and the word cue is used to determine which one will be the target (see Folk & Remington, 2008, and Leonard & Egeth, 2008, for similar results). Thus, it is possible that the present study under-estimates the importance of top-down tuning or over-estimates the influence of the trial history by using the standard protocol with a single (definitive) pop-out or singleton target and word cues that could be ignored (see also Cochrane & Pratt, 2020).

On the other hand, it is however also plausible that feature priming effects could have a strong impact on attention. Both top-down tuning and feature priming are thought to be mediated by a feature weighting mechanism that can dynamically adjust the feature gains, viz., the neuronal response of feature-specific neurons, such that they respond either more or less vigorously in response to a matching stimulus input. The trial history probably affects these gain settings by simply remaining in place: After selection of the target, neurons responding to the target feature remain in an enhanced state, whereas those responding to a nontarget feature may remain in a suppressed state. Once new information about target features is provided (e.g., via a word cue), these feature settings either need to be adjusted (if the target is different from the previous one) or they can remain in place (if the target remains the same). Feature priming could have a larger effect on attention than top-down tuning because it occurs automatically, without fail, with the result that attention is always (at first) biased to the previous target feature (e.g., Becker & Horstmann, 2009).

In turn, top-down tuning is an active process that can fluctuate across trials (e.g., Büsel, Pomper, & Ansorge, 2019; Leber, 2010). Temporal lapses or fluctuations in top-down control would lead to a failure to adjust the feature gain settings for the upcoming target, which would result in a failure to modulate feature priming effects (e.g., Landau & Fries, 2012; Leber, 2010; see also Ansorge & Becker, 2012). Thus, temporal lapses in top-down tuning could explain why top-down tuning did not completely eliminate automatic feature priming effects, or more generally, why top-down tuning may have a weaker effect than automatic priming effects. While this explanation would warrant further research, the present results clearly showed that dynamic feature weighting processes override and govern feature contrast effects.

In particular, the data suggest that increasing the feature contrast only benefited target selection when attention was biased to the nontargets. Sceptics may argue that the data do not directly support nontarget biasing as the driving factor, as it is also possible that feature contrast effects were only neutralised when attention was biased towards the target (in virtue of priming or top-down knowledge). That is, a bias for the target may lead to immediate selection of the target (i.e., a ceiling effect), which eliminates the feature contrast effect that would otherwise benefit search.

This alternative explanation has not been proposed in the literature and would still struggle to explain the effects. First, the data reliably indicate flat set size functions in attentional guidance measures even in the absence of ceiling effects: For instance, the switch trials in the validly cued condition exhibit clear priming effects (switch costs), but no inverse set size effects (see Fig. 3, top graphs). Second, the alternative explanation is inconsistent with the positive set size effects in a subset of our conditions. If an increase in feature contrast usually benefits search, how can we explain that it has the opposite effect when attention is biased to the target? A positive set size effect means that the nontarget items compete for attention with the target as individual items, but, here, this only occurs when attention is biased to the target. This is contrary to the original hypothesis that increasing the number of nontargets helps with target localisation. Hence, if we assume that feature contrast generally benefits selection, we would have to introduce a new process that can change the effects of feature contrast depending on how attention is tuned to the target. This does not seem to be a promising avenue for theory development.

By contrast, the nontarget biasing account can explain the different effects of feature contrast using only known and well-established effects. By assuming that grouping is a late process that automatically happens when an item is (covertly) selected, we can directly derive the prediction of inverse set size effects in all conditions in which attention is biased towards the nontargets (i.e., in which we would observe frequent selection of the nontargets). Grouping cannot take effect when attention is biased to the target because the target is a singleton and there are no other items in the display sharing the target-defining feature. Hence, *if* the bias to the target is weak or if it fails, the target has to initially compete against the (ungrouped) nontargets as individual items, which can account for the finding of positive set size effects in a subset of conditions.

In conclusion, the nontarget biasing account seems to provide a more consistent and parsimonious explanation of the results. The nontarget biasing account can also explain previous discrepancies in the literature. Specifically, it can explain why, in Experiment 2 and in Becker and Ansorge (2013), we failed to find strong feature contrast effects when only the target feature changed or when only the nontargets changed, and why we only obtained clear evidence for feature contrast effects when the target and nontarget colour fully swapped, or when there was a partial swap in the colours (Exp. 1 and 2; see also McPeek et al., 1999).

8. Implications

The present findings are inconsistent with the view that low local feature contrasts impose bottom-up limitations on target selection and shed doubt on the classical view that such bottom-up saliency signals reliably guide attention to the target. The finding that sparse displays do not automatically limit target selection is important because it marks the important distinction between feature contrast (e.g., colour contrast) and local feature contrasts (e.g., display density).

A lack of feature contrast and especially, feature contrasts below the pre-attentive just noticeable difference, reliably limits target selection and leads to inefficient search (i.e., positive set size effects; e.g., Becker & Ansorge, 2013; see also Duncan & Humphreys, 1989; Treisman & Sato, 1990; Wolfe, 1994). However, the strength of a feature contrast in feature space is defined independently of the strength of its local realization in the visual field, and the present results clearly show that the two concepts should not be confused, as what is true of feature contrast is not true for a lack of *local* feature contrasts. Yet, it is still widely believed that using sparse displays and, thus, limiting local feature contrast impairs selection based on feature contrasts in feature space and leads to a qualitatively different search (e.g., Rangelov et al., 2017; see also Meeter & Olivers, 2006). Similarly, theories of attention seem to compute bottom-up saliency mostly on the basis of *local* feature contrasts, often within a spatial region approximating the (enlarged) receptive fields of neurons in V1 (e.g., Itti, Koch, & Niebur, 1998; Navalpakkam & Itti, 2006; Wolfe, 1994).[4] Conversely, our results demonstrate that a lack of local feature contrast does not reliably limit selection and, thus, behaves differently to a lack of feature contrast (e.g., when the target is too similar to the nontargets; Duncan & Humphreys, 1989), which may guide future development of theories of attention.

Moreover, inverse set size effects are only included in a few models of visual search and attention (e.g., Buetti et al., 2016), and are still regarded as somewhat of an ‘anomaly’ (cf. Rangelov et al., 2017). Contrary to the mainstream view that local feature contrasts aid attentional guidance, we found that increased local feature contrasts reliably facilitate decision-level processes involved in nontarget rejection or target identification, as reflected in late measures (i.e., dwell times; Becker & Ansorge, 2013). These findings are at odds with the view that bottom-up feature contrasts have only transient, short-lived effects on visual selection (e.g., Donk & Soesman, 2010; Donk & Van Zoest, 2008) and are in line with studies showing that bottom-up feature contrast has longer-lasting effects on decision-making and memory-related processes (e.g., Constant & Liesefeld, 2021; Kiss, Grubert, Peterson, & Eimer, 2012; Martin & Becker, 2018).

In addition, our finding that bottom-up feature contrasts can aid decision-level processes indicates that target identification does not always merely consist in comparing the feature of the selected item to a target representation (e.g., Duncan & Humphreys, 1989). Rather, target verification in pop-out search seems to involve context-dependent mechanisms that probe the uniqueness of the selected feature by comparing it to the features of the other items in the surround (perhaps by selecting multiple items in parallel; e.g., Venini et al., 2014). This target verification process seems to profit from a larger number of nontargets in the vicinity, explaining the inverse set size effect in the dwell times. Previous studies have already pointed out that varying the set size may also influence the amount of noise in decision-level processes (e.g., Palmer, 1995). However, no extant theory of attention includes a target identification or nontarget rejection mechanism that would explain inverse set size effects.

Thus, our findings can guide future developments in theories, to capture more accurately how local feature contrasts should be included in models, both with respect to attentional guidance and later, decision-making processes.

In addition, our findings revealed rather large differences between early versus late-stage processes in visual search, which were found to be differentially sensitive to the three different manipulations (top-down knowledge, local feature contrast, and trial history; see Fig. 7). These findings support the view that early, attention-guiding processes operate on different mechanisms and require a different explanation than later, decisional processes (Hamblin-Frohman & Becker, 2021; Yu, Hanks, & Geng, 2022; Yu, Zhou, Becker, Boettcher, & Geng, n.d.).

In the present paradigm (pop-out search), we found that feature priming modulated early and intermediate processes prior to selecting the target, whereas it did not affect later processes, indexed by dwell times on the items after selection. This is in line with the view that priming transiently affects early processes of attentional guidance, and had negligible or no effects on later, decision-making processes (e.g., Wolfe, 2021; see also Becker, 2008a, Becker, 2008b, Becker, 2010). Similarly, top-down knowledge about the target mainly exerted its effect at an early and intermediate stage of visual search, but did not modulate later processes as indexed by the dwell times (see Fig. 7).

This large discrepancy across early versus late measures argues against the common approach of using mean RTs or the RT set size function for inferences about early, attention-guiding processes (e.g., Meeter & Olivers, 2006; Rangelov et al., 2017; Wolfe, 1994, Wolfe, 2021). In the present study, we were able to show how using the mean RTs to make inferences about early, attention-guiding processes led to the wrong conclusion that local feature contrasts always guide attention, where in fact, local feature contrast effects mainly resided in the late measures (e.g., dwell time; see also Becker, 2010; Hamblin-Frohman & Becker, 2021; Martin & Becker, 2018, Yu et al., 2022). To accurately identify the source of local feature contrasts or inverse set size effects, future research cannot rely on mean RTs or the slope of the RT set size function, but needs to use measures that tap into early, attention-guiding processes in visual search.

Author note

This research was supported by Australian Research Council (ARC) grant DP210103430 to S. I. Becker. All materials and data of Experiments 1 and 2 (including the stimuli, experiment files and de-identified individual data files) are publicly available via this link: https://www.dropbox.com/sh/fefvjl8nt5btv7p/AACQJKyCL59yb_lc2PtkO1ZCa?dl=0.

Credit author statement

SIB designed and conducted the study, and provided a first draft of the results. GH, AKG and UA provided comments and edits on the manuscript and the response letters.

Data availability

All materials and data are publicly available via a link included in the paper.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*, 437–443.

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

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

Becker, S. I. (2008b). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research*, *48*, 664-684.

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.

Becker, S. I. (2011). Determinants of dwell time in visual search: Similarity or perceptual difficulty? *PLOS ONE*, *6*, Article e17740.

Becker, S. I., & Ansorge, U. (2013). Higher set sizes in pop-out search displays do not eliminate priming or enhance target selection. *Vision Research, 81*, 18-28.

Becker, S. I., & Horstmann, G. (2009). A feature weighting account of priming in conjunction search. *Attention, Perception, & Psychophysics, 71*, 258-272.

Becker, S. I., Lewis, A. J., & Axtens, J. E. (2017). Top-down knowledge modulates onset capture in a feedforward manner. *Psychonomic Bulletin & Review*, *24*, 436-446.

Becker, S. I., Martin, A. & Hamblin-Frohman, Z. (2019). Target templates in singleton search vs. feature-based search modes. *Visual Cognition,* 27*,* 502-517

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

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science, 308*, 529–534.

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics, 51*, 465-472.

Büsel, C., Voracek, M., & Ansorge, U. (2020). A meta-analysis of contingent-capture effects. *Psychological Research*, *84*(3), 784-809.

Büsel, C., Pomper, U., & Ansorge, U. (2019). Capture of attention by target-similar cues during dual-color search reflects reactive control among top-down selected attentional control settings. *Psychonomic Bulletin & Review*, *26*(2), 531-537.

Buetti, S., Cronin, D. A., Madison, A. M., Wang, Z. & Lleras, A. (2016). Towards a better understanding of parallel visual processing in human vision: Evidence for exhaustive analysis of visual information. *Journal of Experimental Psychology: General, 145*, 672-707.

Cochrane, B. A., & Pratt, J. (2020). Re-examining Maljkovic and Nakayama (1994): Conscious expectancy does affect the Priming of Pop-out effect. *Attention, Perception, & Psychophysics, 82*, 2693-2702.

Conway, B. R., Hubel, D. H., & Livingstone, M. S. (2002). Color contrast in macaque V1. *Cerebral Cortex, 12*, 915-925.

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

Eimer, M., & Grubert, A.K. (2014). Spatial attention can be allocated rapidly and in parallel to new visual objects. *Current Biology*, *24*, 193-198.

Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. Jour*nal of Cognitive Neuroscience, 20*(8), 1423-1433.

Eimer, M., Kiss, M., Press, C., & Sauter, D. (2009). The roles of feature-specific task set and bottom-up salience in attentional capture: An ERP study. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1316-1328.

Eriksen, C.W. & St James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*, 225-240.

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

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 847-858.

Folk, C. L., & Remington, R. W. (2008). Bottom-up priming of top-down attentional control settings. *Visual Cognition, 16*, 215-231.

Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences, 22*(1), 79-92.

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.

Helmholtz, H. v. (1867). *Handbuch der physiologischen Optik* [Handbook of physiological optics]. Voss.

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

Horstmann, G., Becker, S. I., & Grubert, A. (2020). Dwelling on simple stimuli in visual search. *Attention, Perception, & Psychophysics, 82,* 607-625.

Huang, L., Holcombe, A. O., & Pashler, H. (2004). Repetition priming in visual search: Episodic retrieval, not feature priming. *Memory & Cognition*, *32*(1), 12-20.

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review, 114*(3), 599–631.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489-1506.

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

James, W. (1890). *The principles of psychology*. Holt.

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.

Koch, C., & Ullman, S. (1985). Shifts in visual selective attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*, 219-227.

Kristjánsson, Á., & Driver, J. (2008). Priming in visual search: Separating the effects of target repetition, distractor repetition and role reversal. *Vision Research, 48*, 1217-1232.

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.

Lamy, D., Darnell, M., Levi, A., & Bublil, C. (2018). Testing the attentional dwelling hypothesis of attentional capture. *Journal of Cognition*, *1*(1), Article 43.

Leber, A. (2010). Neural predictors of within-subject fluctuations in attentional control. *Journal of Neuroscience, 30*, 11458-11465.

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.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6*, 9–16.

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress towards resolving the attentional capture debate. *Visual Cognition, 29*(1), 1-21.

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

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*, 1865-1883.

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.

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Cognitive Sciences, 29*, 317-323.

McDonald, J. J., Green, J. J., Jannati, A., & Di Lollo, V. (2013). On the electrophysiological evidence for the capture of visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 849–860.

McPeek, R.M., & Keller, E.L. (2002). Superior colliculus activity related to concurrent processing of saccade goals in a visual search task. *Journal of Neurophysiology*, *87*, 1805-1815.

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.

McPeek, R.M., Skavenski, A.A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, *40*, 2499-2516.

Meeter, M., & Olivers, C. N. L. (2006). Intertrial priming stemming from ambiguity: A new account of priming in visual search. *Visual Cognition, 13*, 1-28.

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

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research, 51*, 1526-1537.

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America, 7*(7), 1209-1217.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387-391.

Navalpakkam, V., & Itti, L. (2006). An integrated model of top-down and bottom-up attention for optimizing detection speed. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2049-2056).

Navalpakkam, V., & Itti, L. (2007). Search goals tunes visual features optimally. *Neuron, 53*, 605-617.

Rangelov, D., Müller, H. J., & Zehetleitner, M. (2017). Failure to pop out: Feature singletons do not capture attention under low signal-to-noise ratio conditions. *Journal of Experimental Psychology: General, 146*, 651-671.

Reynolds, J. H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron, 37*(5), 853-863.

Shiu, L.-P., & Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1037–1054.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261-267.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science, 240*, 338-340.

Sprague, T. C., Itthipuripat, S., Vo, V. A., & Serences, J. T. (2018). Dissociable signatures of visual salience and behavioral relevance across attentional priority maps in human cortex. *Journal of Neurophysiology, 119*(6), 2153-2165.

Theeuwes, J. (2013). Feature-based attention: it is all bottom-up priming. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368,* 20130055.

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

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*, 65-70.

Theeuwes, J., Kramer, A.F., Hahn, S., & Irwin, D.E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Human Factors*, *9*, 379-385.

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

Van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746–759.

Venini, D., Remington, R.W., Horstmann, G., & Becker, S.I. (2014). Centre-of-gravity fixations in visual search: When looking at nothing helps to find something. *Journal of Ophthalmology, 237812*, 1-14.

Walker, R. & McSorley, E. (2006). The parallel programming of voluntary and reflexive saccades, *Vision Research*, *46*, 2082-2093.

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

White, B. J., Kan, J. Y., Levy, R., Itti, L., & Munoz, D. P. (2017). Superior colliculus encodes visual saliency before the primary visual cortex*. Proceedings of the National Academy of Sciences of the United States of America*, *114*(35), 9451-9456.

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

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review, 28*, 1060-1092.

Yu, X., Hanks, T.D., & Geng, J. (2022). Attentional guidance and match decisions rely on different template information during visual search. *Psychological Science*, *33*, 105-120.

To cite this article: Becker, S. I., Grubert, A., Horstmann, G., & Ansorge, U. (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