



Concurrent evaluation of independently cued features during perceptual decisions and saccadic targeting in visual search

Doug J. K. Barrett¹ · Oliver Zobay²

Published online: 9 September 2019

© The Author(s) 2019

Abstract

Simultaneous search for one of two targets is slower and less accurate than search for a single target. Within the Signal Detection Theoretic (SDT) framework, this can be attributed to the division of resources during the comparison of visual input against independently cued targets. The current study used one or two cues to elicit single- and dual-target searches for orientation targets among similar and dissimilar distractors. In Experiment 1, the accuracy of target discrimination in brief displays was compared at set sizes of 1, 2 and 4. Results revealed a reduction in accuracy that scaled with the product of set size and the number of cued targets. In Experiment 2, the accuracy and latency of observers' saccadic targeting were compared. Fixations on single-target searches were highly selective towards the target. On dual-target searches, the requirement to detect one of two targets produced a significant reduction in target fixations and equivalent rates of fixations to distractors with opposite orientations. For most observers, the dual-target cost was predicted by an SDT model that simulated increases in decision-noise and the distribution of capacity-limited resources during the comparison of selected input against independently cued targets. For others, search accuracy was consistent with a single-item limit on perceptual decisions and saccadic targeting during search. These findings support a flexible account of the dual-target cost based on different strategies to resolve competition between independently cued targets.

Keywords Visual search · Selection · Decision · Capacity

Introduction

Finding a target-object in the scene requires observers to compare visual input at different locations with an internal representation of the target's features. This ability is thought to rely on components of selective attention that integrate information across frontal, parietal and visual cortical regions (Ptak, 2012). Prior to search, an attentional template in visual short-term memory maintains a description of the target's visual features (i.e., its colour and orientation or categorical identity: Carlisle, Arita, Pardo, & Woodman, 2011; Duncan & Humphreys, 1989; Desimone & Duncan, 1995; Olivers, Peters, Houtkamp, & Roelfsema, 2011). During search,

information in the attentional template is used to weight relevant over irrelevant visual input across the visual field (Chelazzi, Duncan, Miller, & Desimone, 1998; Martinez-Trujillo & Treue, 2004; Maunsell & Treue, 2006). This top-down control-signal mediates interactions in topographical cortices that represent the locations of target features (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Ester, Sutterer, Serences, & Awh, 2016; Ling, Liu, & Carrasco, 2009; Müller et al., 2006). Within a single fixation, selective attention protects capacity-limited decision processes by restricting the evaluation of visual input to locations containing task-relevant features (Bundesen, Habekost, & Kyllingsbæk, 2005; Eimer, 2015; Folk, Remington, & Wright, 1994; Wolfe & Gray, 2007). When search operates over longer durations, selection also informs the programming of saccades to locations that are most likely to contain the target (Bisley & Goldberg, 2010; Rutishauser & Koch, 2007; Zelinsky & Bisley, 2015). Within this framework, the accuracy of perceptual decisions and saccadic sampling depend on the specificity of the information that informs selection, as well as the decision processes that operate on objects at selected locations (Eimer, 2015; Palmer, 1995).

Most studies of search have focussed on the accuracy and speed of detection for a single target. In this situation,

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13414-019-01854-w>) contains supplementary material, which is available to authorized users.

✉ Doug J. K. Barrett
djkb1@leicester.ac.uk

¹ Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester LE1 9HN, UK

² Hearing Sciences, University of Nottingham, Nottingham, UK

selection elicits a topographic map that represents objects in terms of their similarity to the target's features (Zelinsky & Bisley, 2015). In many situations, however, observers are required to detect more than one possible target. Airport baggage screeners, for example, search for different types of objects (e.g. guns, knives or aerosols). Previous research has shown that this type of search is difficult: Search for one of two targets is often slower and less accurate than separate searches for the same two targets. This 'dual-target' cost has been reported for targets that differ from non-targets (distractors) on a value within a single feature dimension (i.e., colour), as well as a conjunction of values from different feature dimensions (i.e., colour and orientation; Kaplan & Carvellas, 1965; Menneer, Barrett, Phillips, Donnelly, & Cave, 2007; Menneer, Cave, & Donnelly, 2009; Wing & Allport, 1972). The dual-target cost has been reported in brief displays that preclude eye movements (Houtkamp & Roelfsema, 2009) and free-view displays, where the accuracy of saccadic sampling appears to be reduced compared to single-target searches (Dombrowe, Donk, & Olivers, 2011; Stroud, Menneer, Cave, & Donnelly, 2012). Whether these findings reflect a limit on the number of attentional templates that can inform selection at any one time, or post-selection limits on the evaluation of selected objects against multiple decision criteria, is currently unknown.

In the studies above, the dual-target cost describes a reduction in performance associated with the requirement to detect one of two possible targets. In this context, 'single-' and 'dual-target search' denote the number of cued values the observer must evaluate to detect the presence or absence of a single target. As such, the terms reflect changes to the observer's attentional set rather than the number of targets that can appear in the display (Barrett & Zobay, 2014; Cave, Menneer, Nomani, Stroud, & Donnelly, 2018; Irons, Folk & Remington, 2012; Stroud et al., 2012). One explanation for the dual-target cost is that selection during search is limited to a single-item attentional template (SIT: Beck & Hollingworth, 2017). Evidence consistent with this account was reported by Houtkamp and Roelfsema (2009), who used a rapid serial presentation (RSVP) task to compare detection accuracy for briefly presented targets preceded by one or two cues. Their data revealed a reliable decrease in detection accuracy for targets preceded by two cues, which was best accounted for by a signal-detection (SDT) model that limited search to a single target. Support for a SIT limit on selection has also been obtained by studies investigating the effects of visual short-term memory (VSTM) load on search. For example, van Moorselaar and colleagues (van Moorselaar, Theeuwes, & Olivers, 2014) required observers to remember a variable number of colours prior to onset of search display. At a VSTM load of one, their data revealed a significant increase in response times (RTs) when the colour of the memory item matched that of a distractor during search. When memory

displays contained more than one remembered colour, however, search RTs were unaffected (see Hollingworth & Beck, 2016, for contradictory evidence). Based on these and similar results, Olivers and colleagues (Olivers et al., 2011) proposed a functional distinction between objects maintained in VSTM. During search, the attentional template has direct access to the mechanisms that select relevant visual input. Accessory objects, which are outside the current focus of attention, are maintained in a passive state that neither contributes to nor interferes with the selection of visual input (Downing & Dodds, 2004; Olivers & Eimer, 2011).

The SIT model imposes a bottleneck between feature-based information in the attentional template and visual input during search. In brief displays, this predicts selection that is exclusive to a single target (Moore & Weissman, 2011, 2014). In longer displays, the bottleneck should slow search, because observers must switch the status of objects in VSTM to select and evaluate the presence of different targets. Recent behavioural and electrophysiological data, however, indicate pre-cues can modulate visual input for more than one target at the same time. Grubert and Eimer (2016) used event-related potentials (ERPs) to compare N2pc components to coloured targets that were preceded by coloured cues. On single-target trials, one of two pre-cued colours was presented in the search display. On dual-target trials, both pre-cued colours were presented in quick succession or simultaneously. The N2pc is thought to index the spatial selection of goal-relevant features in the scene (Cohen, Heitz, Schall, & Woodman, 2009; Eimer, 1996, 2015; Luck & Hillyard, 1994), and the data revealed comparable amplitudes and latencies for sequential and simultaneously presented colour targets preceded by one or two cues.

The electrophysiological results above are consistent with the parallel selection of visual input by an attentional template that encodes multiple target features (multiple-item template; MIT). Evidence to suggest independently cued features can also mediate perceptual categorisation was reported by Roper and Vecera (2012). They used a cued-RSVP task to elicit single- and dual-target searches for centrally presented objects in the presence of peripheral distractors. Distractors that matched the colour of one or two pre-cues decreased the accuracy of target detection; indicating attentional capture for different colours within a single frame (50 ms). Coloured distractors that matched neither cue had no effect on the accuracy of target detection, ruling out a stimulus-driven explanation for attentional capture in their task. Multiple colour cues have also been shown to mediate attentional capture during single-target search. Irons and colleagues contrasted RTs for targets that were presented at the same or a different location to a spatial cue that was red, green or blue. Prior to the task, red and green were specified as target colours and the data revealed spatial validity effects for targets at locations that were

cued by either target colour (Irons, Folk, & Remington, 2012). This result is consistent with top-down modulation of visual input at the cued location for independently cued features. Notably, however, RTs tended to be faster for congruent compared to incongruent cue-target pairs, illustrating an asymmetry in the facilitation afforded by the two cued colours. Irons and colleagues interpreted this as evidence of a dissociation between selection and post-selection processes; with incongruent cue-target pairs disrupting decision processes or the consolidation of targets in VSTM, rather than the selection of task-relevant visual input (see also Adamo, Wozny, Pratt, & Ferber, 2010).

The results above suggest the selection of visual input for independently cued targets generalises from electrophysiological to behavioural responses during search. Evidence that MITs can also guide oculomotor behaviour has been reported by Beck and colleagues (Beck, Hollingworth, & Luck, 2012). They required observers to search multi-coloured arrays for an orientation-defined target. Single- and dual-target searches were elicited with one or two cues and observers were instructed to conduct sequential or simultaneous searches for both colours on dual-target trials. The data revealed sequences of fixations to objects of the same or alternate cued-colours proceeded at a similar rate, indicating observers could switch templates between saccades without incurring a delay. A comparison of the proportion of fixations to cued and uncued colours also revealed equivalent selectivity on single- and dual-target searches. Despite comparable accuracy, however, both fixation duration and manual responses were slower for dual- than single-target searches. Using a different method, Barrett and Zobay (2014) contrasted the speed and accuracy of single- and dual-target searches in free-view search displays. Their data revealed comparable slopes, but an increase in the intercepts of set size-by-RT functions for dual-target searches. SDT estimates fitted to the proportions of correct responses on target-present and -absent trials were also consistent with a small decrease in target discriminability when two targets were cued. The results of both studies indicate observers can use concurrently active features to guide search, but that the requirement to do so incurs costs in terms of the speed and accuracy of saccadic sampling and perceptual decisions.

The findings above are consistent with top-down selection of visual input by a multiple-item template during search. The distinction between ERP evidence for equivalent target-modulation and behavioural evidence of reductions in the speed and accuracy of target-detection, however, suggests the benefits of selection may not generalise to the mechanisms that support perceptual judgements and/or saccadic guidance during dual-target search. One possibility is that the single-item limit observed in previous studies reflects a constraint on

the number of decision criteria that can be applied to selected objects during search. In this case, evidence for a single-item limit on attentional control would reflect a constraint on post-selection resources during dual-target search. Competition between selected objects based on similarity criteria for different cues may also disrupt the planning and execution of saccades during search (Godijn & Theeuwes, 2002). In this case, evidence for a single-item template may reflect the strategic prioritisation of a single-decision process during the programming of saccades to independently cued features (Cave et al., 2018; Grubert & Eimer, 2015; Rajsic, Ouslis, Wilson, & Pratt, 2017). Alternatively, dual-target costs in behavioural studies may reflect the distribution of post-selection resources across similarity distributions for separately cued targets (Barrett & Zobay, 2014; Menneer et al., 2009). In the absence of capacity-constraints, searching for one of two targets will increase decision-noise, because detection entails the comparison of visual input with different decision criteria. In this situation, decisions based on independent similarity distributions would increase the joint probability of distractors being confused with either target (Eckstein, 1998; Eckstein, Thomas, Palmer, & Shimozaki, 2000). The requirement to distribute resources to compare visual input against multiple cues may also reduce the accuracy of decision processes. Data from change-detection studies have shown that the precision of recall is inversely related to the number of remembered objects (Bays & Husain, 2008; Burmester & Wallis, 2012; Salmela & Saarinen, 2013), and recent evidence suggests this may reflect changes in the rate at which selected information is transferred into VSTM (Sewell, Lilburn, & Smith, 2014; Smith & Sewell, 2013). In dual-target search, selecting input associated with separately cued features may slow the transfer of selected information in VSTM, decreasing the fidelity of perceptual categorisation in brief displays (e.g. Irons et al., 2012) and slowing saccades during overt search (e.g. Beck et al., 2012).

The purpose of the current study was to investigate the nature of the dual-target cost during covert and overt search. Experiment 1 investigated the accuracy of perceptual judgements in brief displays that precluded eye movements on single- and dual-target searches. Experiments 2 investigated the accuracy of initial saccadic targeting (Awh, Armstrong, & Moore, 2006) on single- and dual-target searches. In the following experiments, we generalise the SDT models used by Houtkamp and Roelfsema (2009) to assess whether changes in target discriminability on dual-target searches are consistent with those predicted by (i) a single-item template limit on selection during search or (ii) multiple-item templates when two objects are cued. In the former, the dual-target cost is modelled by restricting target-present or -absent decisions to a single-item attentional template on dual-target searches (i.e., one cued orientation). In the latter, the dual-target cost is modelled as a function of the product

of set size and the number of templates guiding search (i.e., 1 or 2 * set size on single- and dual-target searches, respectively). In addition, we derive noise- and capacity-limited versions of the single- and multiple-item template models to distinguish dual-target costs associated with decision-noise from those associated with a decrease in target discriminability when observers search for one of two cued targets (Barrett & Zobay, 2014; Houtkamp & Roelfsema, 2009; Smith, Lilburn, Corbett, Sewell, & Kyllingsbæk, 2016). Crossing single- and multiple-item templates with noise- and capacity-limited derivations produces four models, which are used to evaluate the contributions of decision-noise and capacity constraints on the dual-target costs predicted by distinct accounts of attentional control during dual-target search. To our knowledge, this is the first study to contrast the predictions of different explanatory models of the dual-target cost on the accuracy of perceptual decisions during covert search and saccadic targeting in free-view displays.

Experiment 1

Method

Observers

We used a small-*n* design to evaluate the descriptive and explanatory adequacy of SDT models of search. Our focus, therefore, is on the relationship between the observed and estimated data at individual and group levels of analyses. To equate sample size with relevant EEG studies (i.e., Eimer & Grubert, 2014; Grubert & Eimer, 2015) we recruited 12 observers. Data from one participant were excluded because they withdrew before completing the study. Of the remaining sample, seven were male and their ages ranged from 20 to 46 years ($M_{\text{age}} = 23.20$, $SD = 4.96$). All reported normal or corrected-to-normal visual acuity. Recruitment, consent and all experimental procedures conformed to American Psychological Association (APA) ethics standards.

Apparatus

The experiment was run on an IBM PC with a 19-in. CRT View Sonic G90fB monitor (Walnut, CA, USA). The display resolution was 1,240 × 768 pixels and the frame rate was 85 Hz. Stimulus presentation and data collection were controlled using custom-built software in MATLAB (Mathworks, Natick, MA, USA) with Psychophysics toolbox extensions (Brainard & Vision, 1997; Kleiner et al., 2007). Viewing distance was maintained at 57 cm using a fixed chin rest and responses were collected using a Cedrus RB-350 Response

Pad (San Pedro, CA, USA). The experiment was conducted in a quiet, dimly lit room.

Stimuli

Displays contained red ($x = 0.590$, $y = 0.326$, $Y = 4.54$ cd/m²) Landholt's *C*-shapes and grey ($x = 0.281$, $y = 0.303$, $Y = 4.54$ cd/m²) annuli that subtended 3.0° × 3.0°. Stimuli were presented at 12 equally spaced locations on the circumference of a virtual circle with a radius of 7.0°. *C*-shapes and annuli were presented on a uniform black ($x = 0.590$, $y = 0.326$, $Y = 0.55$ cd/m²) background.

Procedure

The experiment used a factorial design to manipulate four independent variables: Search type (single- or dual-target); trial type (target-present or -absent); target identity (Target-Left or -Right), and set size (1, 2 or 4 *C*-shapes). Experimental blocks contained four repetitions of this structure (96 trials). On each block of trials, two target and four distractor orientation values were assigned. Target orientations were sampled from 90° and 270° ± 15° to 30°. Distractor orientations for each observer differed from Target-Left or Target-Right by an angle of rotation that produced 80% accuracy on a pre-test (see below). Target and distractor values were fixed within experimental blocks to produce symmetric target-distractor similarity distributions for two targets among numerically equivalent subsets of leftward and rightward *C*s. (i.e., ± ~180°, see Fig. 1). This symmetry was designed to control potential differences in the perceptual salience of cued objects that might bias observers towards a serial strategy favouring one target over the other. The use of leftward and rightward *C*-Shapes was also designed to elicit competing similarity values between independently cued targets and numerically identical subsets of objects on dual-target searches. Displays could contain 1, 2 or 4 *C*-shapes, with the remaining locations occupied by an annulus to control for differences in perceptual crowding across set size.

Figure 2 illustrates the sequence of events on each trial. Trials began with a fixation cross at the centre of the screen. After 500 ms, the fixation was replaced by a cue containing two objects centred on the horizontal midline ± 3° from fixation. Cues signalled the orientation of the subsequent search-target. On single-target trials, the cue consisted of an annulus and *C*-shape at the orientation assigned to Target-Left or Target-Right. On dual-target trials, the cue contained two *C*-shapes – one at the orientation assigned to Target-Left and one at the orientation assigned to Target-Right. To equate inspection times, single- and dual-target cues were presented for 494 ms and 988 ms, respectively. Cues were followed by a 988-ms blank inter-stimulus-interval before the onset of the search display. At a set size of four, target-absent displays

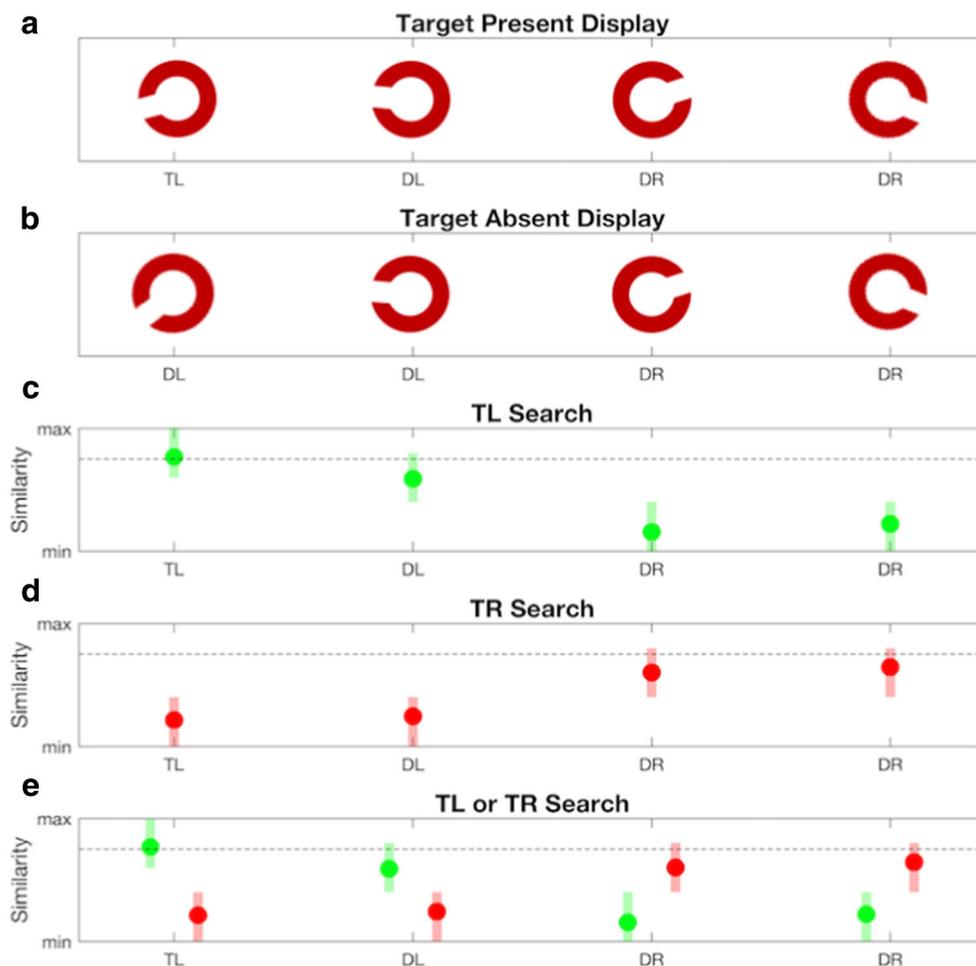


Fig. 1 (a) Illustration of the orientation values assigned to C-shapes in a search display containing a leftward-leaning (TL) target. (b) Illustration of the orientation values assigned to C-shapes on a target-absent display. Distractors with similar and dissimilar orientations to TL are labelled DL and DR, respectively. (c) Hypothetical similarity values (filled circles) drawn from a normal distribution of internal responses (rectangles) to the C-shapes in panel A during a single-target search for TL. (d)

Hypothetical similarity values and internal responses to the C-shapes in panel B during search for an absent rightward-leaning target. (e) Hypothetical similarity values and internal responses to the C-shapes in panel A during a dual-target search where observers are cued to search for a rightward- or leftward-leaning target. Dotted horizontal lines depict response criteria for target-present responses (similarity > criteria = target)

contained one C-shape at the orientation assigned to each of the distractors (i.e., TL:R $\pm \Delta^\circ$). At set size two, one leftward and one rightward distractor were randomly sampled from the four alternatives. At set size one, a single leftward or rightward distractor was sampled with equal probability. On target-present trials, displays contained one C-shape at the orientation assigned to Target-Left or Target-Right. Targets always replaced a similar distractor from the same (leftward or rightward) group and displays always contained one C-shape from a group that was cued. Search displays were presented for 94 ms and followed by blank screen, which remained visible until a response was recorded. Short-duration displays were used to equate processing time and prevent eye movements on single- and dual-target searches.

Observers completed ten blocks of trials in a single experimental session. In each block, single- and dual-target cues were equally likely for leftward and rightward targets. A cued

target appeared in the search display on 50% of trials, and the order of presentation for each target, search type and trial was randomly assigned for each block. Prior to experimental sessions, observers completed two pre-test blocks of 80 trials. These presented a single C-shape in the same displays as the experimental session but varied the orientations of the target and distractor across the range $\pm 5^\circ$ to 30° using a method of constant stimulus. Individual responses were fitted with a cumulative Gaussian function to estimate the angle of rotation required by each observer to distinguish targets from distractors on 80% of trials.

SDT models of search

The observer's task on each trial is to classify the display in terms of the presence or absence of a target ('yes-no' decision). The statistical modelling of this process is

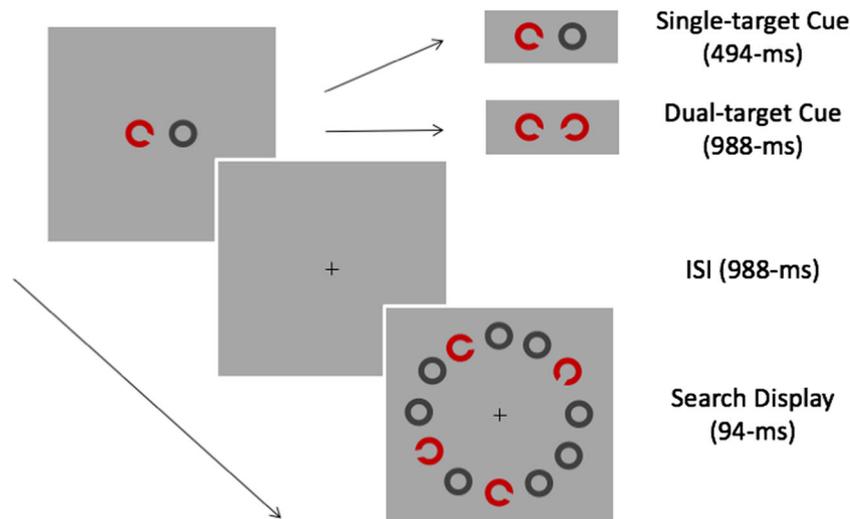


Fig. 2 Sequence of events on each trial of the single-fixation search task. The search display contains the cued target (bottom) as well as a similar (upper left) and two dissimilar distractors

based on signal detection theory (SDT) and the assumption of independent comparisons (Palmer, Verghese, & Pavel, 2000; Shaw, 1982). To classify the display, the observer independently compares the internal representation of orientation for each cue i to the visual response to each object in the display j . The result of each comparison describes the perceived similarity between the cue and the object, which is represented by a real number s_{ij} . To classify the display, observers evaluate whether the maximum similarity value is less than or equal to a response criterion. The cumulative probability distribution function $P(s_{ij} \leq R)$ of s_{ij} depends on three factors: (1) The type of comparison $C(i, j)$, which can be a target (t) similar distractor (s) or dissimilar distractor (d), depending on the relationship between the cue and the object in the display. (2) The set size D , which can equal 1, 2 or 4. (3) The number T of templates used to guide search, which is determined by the number of cues. These dependencies are indicated by the notation $P(s_{ij} \leq R | C, D, T)$. As a simplifying assumption, all comparisons between cues and objects with opposite orientations are modelled by the same probability distribution. In particular, the target is considered a dissimilar distractor for the second cue on dual-target searches. In agreement with standard SDT (Green & Swets, 1974; Wickens, 2002), probability distributions are modelled as Gaussian with mean $\mu(C, D, T)$ and variance 1, i.e., $P(s_{ij} \leq R | C, D, T) = \Phi(R - \mu(C, D, T))$, where Φ denotes the standard normal cumulative probability distribution (Fig 3).

To distinguish between theoretical accounts of attentional control during search, we introduce additional assumptions to derive two types of SDT model. Multiple-item template (MIT) models simulate display classification based on the evaluation of similarity values for independently cued feature values

during dual-target search. We assume that the observer reports target-absent if none of the comparisons exceed a response criterion $\lambda(D, T)$. Because of independence, the probability that none of the comparisons exceed the response criterion equals the product of the probabilities for each individual comparison being less than $\lambda(D, T)$:

$$P(\text{target-absent}) = \prod_{\{i,j\}} P(s_{ij} \leq \lambda(D, T) | C(i, j), D, T) \quad (1)$$

In Eq. 1, the product is over all pairwise comparisons between cues and objects at a given set size and type of search (D, T). Conditional probabilities for hits and false alarms are obtained as $1 - P(\text{target-absent})$, depending on whether the target is present or absent in the display.

Applying Eq. 1 directly to the observed data would require the estimation of six decision thresholds $\lambda(D, T)$ and 18 $\mu(C, D, T)$ parameters (i.e., separate decision thresholds and μ estimates for single- and dual-target searches by comparison type and set size). However, the experimental design yields only 12 observations for each subject (i.e., the hit and false-alarm rates in each of the search types by set size conditions). To derive useful and testable models, we therefore impose restrictions on the model parameters: First, we note that model predictions do not change if the same constant value is added to all parameters appearing together in a product term (1). To fix the absolute scale, we therefore set $\mu(C=t, D, T) = 0$ (i.e., the target distribution is centred at zero). Target discriminability is modelled as the difference between the means of the distributions for the target and those for similar and dissimilar distractors. For similar distractors, $d'_S(D, T) = \mu(C=s, D, T) - \mu(C=t, D, T)$. For dissimilar distractors, $d'_D(D, T) = \mu(C=d, D, T) - \mu(C=t, D, T)$. We then distinguish between noise- and capacity-limited search. In the former, we assume that d' is independent of set size and search type, i.e., $d'(C, D, T) = d'(C)$. As a shorthand, we set $d'(C=s) = d'_S$ and

$d'(C=d) = d'_D$. For capacity-limited models, we assume that d' scales with the inverse root of the total number of comparisons in each search: with $d'(C=t)=0$, this can be written as $d'(C=s, D, T) = d'_S/\sqrt{(D*T)}$ with $d'_S=d'(C=s, D=1, T=1)$. This parameterisation produces negative d' values for distractors that are differentiated from the target distribution on the basis of their dissimilarity from the cue, while scaling d' by $1/\sqrt{(D*T)}$ conforms to the decline in discriminability predicted by the SDT sample-size model of search (see Corbett & Smith, 2017; Palmer, 1994; Smith & Sewell, 2013). For MIT models, we also make an equivalence assumption between searches that entail the same number of comparisons. This implies that single-target searches for $D=2$ or 4 are equivalent to dual-target searches for $D=1$ or 2, respectively. For d' , this assumption is consistent with both noise- and capacity-limited models. For λ , we set $\lambda(D, T) = \lambda(D*T)$, as the number of comparisons is the product of set size and cued orientations. The various λ values are denoted $\lambda_1, \lambda_2, \lambda_4, \lambda_8$. Equation 1 provides a general form that can be applied to estimate the proportion of hits and false alarms across each search type by set size condition. For example, the estimated hit rate for the noise-limited MIT model when Target-Left is present at set size 4 is given by

$$P_{MIT}(\text{Hit}|D = 4, T = 2) = 1 - \Phi(\lambda_8)\Phi(\lambda_8 - d'_S)^3\Phi(\lambda_8 - d'_D)^4 \quad (2)$$

as there is one cue-target comparison (CL-TL), three comparisons with similar distractors (CL-DL & CR-DR * 2) and four comparisons with dissimilar distractors (CR-TL, CR-DL & CL-DR * 2).

In contrast to MIT models, single-item template (SIT) models simulate search guided by a single attentional template. On a dual-target search, target detection is based on pairwise comparisons between the active attentional template and the objects in the display. We assume that the observer reports target absent if these comparisons fall below a response criterion $\lambda(D, T)$. To represent the qualitative distinction between cues during the comparison process, SIT derivations contain separate terms to compute similarity values for objects depending on which template is active during search. For example, the estimated hit rate for the noise-limited SIT model on a dual-target search with a set size of 4 is given by

$$P_{SIT}(\text{Hit}|D = 4, T = 2) = 1 - 0.5 * (\Phi(\lambda_4)\Phi(\lambda_4 - d'_S)\Phi(\lambda_4 - d'_D))^2 + \Phi(\lambda_4 - d'_S)^2\Phi(\lambda_4 - d'_D)^2 \quad (3)$$

In this equation, the observer randomly selects one of the cues and performs a single-target search based on the selected (active) attentional template. If the selected template matches the target in the display, performance will be equivalent to a single-target search. If the selected template matches the target that does not appear, hits are highly unlikely, because the cue-target comparison will yield a

low similarity value. Accuracy across dual-target searches is predicted to be the average of the two terms. As for MIT models, we also distinguish noise- and capacity-limited derivations of SIT search. For the former, we assume d' is independent of set size, i.e., $d'(C, D, T) = d'(C)$. For the capacity-limited model, d' scales with the inverse root of the number of objects in the display, because the number of active attentional templates is fixed at 1. With $d'(C=t)=0$, this can be written as $d'(C=s, D, T) = d'_S/\sqrt{D}$ with $d'_S = d'(C=s, D=1, T=1)$. Similarly, we set $\lambda(D, T) = \lambda(D)$ to model decision-noise and changes in threshold as a function of the number of comparisons between one attentional template and objects in the display. The various λ values are denoted $\lambda_1, \lambda_2, \lambda_4$.

The MIT and SIT models above simulate distinct strategies of attentional control during dual-target search. For MIT models, accuracy on single- and dual-target searches is predicted by an inverse relationship between the product of set size and the number of orientations cued prior to search. For SIT models, accuracy on dual-target searches is predicted by a comparison process that is restricted to a single cued orientation during search (Houtkamp & Roelfsema, 2009). For MIT and SIT models, noise-limited derivations capture an important general property of SDT, which is the predicted association between the probability of false alarms and the number of comparisons required to classify the display: Increasing the number of comparisons during search has a multiplicative impact on the likelihood of mistaking a distractor for a target (Eckstein, 2011; Verghese, 2001; Wilken & Ma, 2004). Capacity-limited derivations of MIT and SIT assume an additional dependency between target discriminability and the number of comparisons required to classify the display.

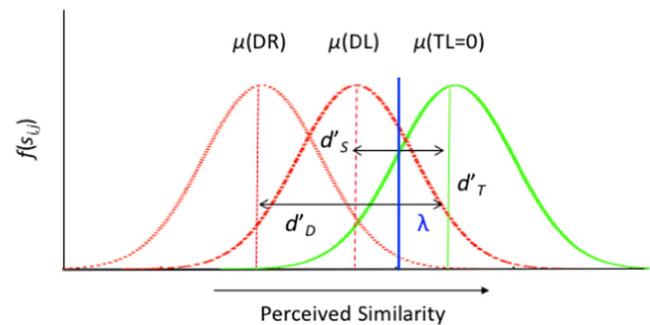


Fig. 3 Illustrates the SDT model applied to a single-target search for a leftward leaning target (TL). The rightmost distribution represents the probability distribution elicited by the target. The centre and leftmost distributions represent the probability distributions elicited by the leftward (DL) and rightward distractors (DR). The target-distribution is centred at zero, and DL and DR distributions are centred at $\mu(DL)$ and $\mu(DR)$, respectively. Target discriminability is modelled in terms of the difference (d') between the means of the distribution elicited by the target and those for similar (i.e., $d'_S = \mu(DL) - \mu(TL)$) and dissimilar (i.e., $d'_D = \mu(DR) - \mu(TL)$) distractors, respectively. The observer responds target present if the perceived similarity of any comparison exceeds the response criterion (λ)

Comparing noise- and capacity-limited versions of the MIT and SIT models provides a means to distinguish the contribution of decision-noise from changes in sensitivity during single- and dual-target searches (Palmer, 1994; Smith & Sewell, 2013). To contrast these predictions, SIT and MIT, noise- and capacity-limited models were fitted to each observer’s data. This required the estimation of five or six parameters (two d' and three or four λ values) producing 5 or 6 error degrees of freedom. Fitting was carried out using maximum likelihood estimation and goodness of fit was assessed using a parametric bootstrap technique for Pearson’s χ^2 statistic. To do this, SDT models were fitted to each observer’s data to compute the observed χ^2 statistic. Parameter estimates for the fitted model were then used to simulate observations over K repetitions of the experiment (we used $K = 500$). By re-fitting and computing the χ^2 statistic for each repetition, the χ^2 distribution that would derive under the fitted model was estimated. Comparing the observed χ^2 statistic against this χ^2 distribution yielded the probability (p -value) that the fitted model generates data that are at least as extreme as those observed. The larger this probability, the ‘more typical’ the observed data are under the fitted model, and the stronger the evidence for its plausibility. As an additional, more intuitive measure of fit, we also calculated the mean absolute difference between corresponding observed and estimated probabilities. For brevity, we report mean group estimates in the text. Model estimates for individuals are reported in the Electronic Supplementary Material (ESM), where the mean absolute difference between observed and estimated values is labelled as ‘Fit 2’.

Results

Orientation thresholds

Best-fitting Gaussian probability density functions for each observer yielded a mean 80% accuracy threshold of $\pm 19.72^\circ$ (SD = 3.1°) at a set size of one.

Accuracy

Accuracy for leftward and rightward leaning targets did not differ significantly (mean 0.703 vs. 0.722; $t(10) = 1.31$, $p > 0.20$, Cohen’s $d = 0.34$), and the remaining analyses collapse responses across both targets during search. Table 1 presents the mean proportion of hits, false alarms and correct responses by set size for single- and dual-target searches. As expected, accuracy was lower on dual- than single-target searches. A 2×3 repeated-measures ANOVA on proportion-correct data yielded significant main effects of search type, $F(1,10) = 144.80$, $p < 0.001$, $\eta_p^2 = .94$, set size, $F(2,20) = 138.80$, $p < 0.001$, $\eta_p^2 = .93$, and a significant Search Type by Set Size

Table 1 Mean proportion of hits [P(H)], false alarms [P(F)] and correct responses [P(C) = (P(H) + 1-P(F))/2] by Search Type and Set Size in Experiment 1

Set size	Single target			Dual target		
	1	2	4	1	2	4
P(H)	.94 (.02)	.80 (.03)	.64 (.03)	.80 (.03)	.63 (.03)	.55 (.05)
P(F)	.24 (.04)	.24 (.03)	.35 (.03)	.22 (.04)	.32 (.03)	.45 (.05)
P(C)	.85 (.02)	.78 (.02)	.65 (.02)	.79 (.02)	.66 (.02)	.55 (.01)

Standard errors in parenthesis

interaction, $F(2,20) = 4.97$, $p = 0.018$, $\eta_p^2 = 0.32$. Two-template SDT models predict a reciprocal relationship between search accuracy and the number of comparisons required to classify the display. Planned contrasts to compare single- and dual-target searches across equivalent set sizes revealed non-significant differences at two, $t(10) = .14$, $p = 0.89$, $d = 0.04$, and four comparisons, $t(10) = .67$, $p = 0.52$, Cohen’s $d = 0.21$, respectively. These data provide initial evidence that accuracy is related to the number of comparisons required to classify the display rather than qualitatively different strategies of attentional control on single- and dual-target searches.

Model estimates and fits

Figure 4 plots the mean observed against the mean predicted proportions of hits and false alarms for each SDT model by search type and set size. Panels A and B reveal predicted values for the SIT models that diverge considerably from the observed data. Panels C and D, in contrast, indicate much better fits between predicted and observed data, with the MIT, capacity-limited model producing mean predicted values within 1 standard error of the observed means for single- and dual-target searches across set sizes. At set size 4, predicted hit and false-alarm rates for the MIT, noise-limited model fall outside the standard error of observed values on single- and dual-target searches, indicating a worse fit than for the MIT, capacity-limited model. Table 2 presents mean parameter estimates of target-discriminability (d') and response bias (λ) across equivalent set sizes, as well as statistical comparisons for the four models. All reveal a monotonic increase in λ as a function of set size. Estimates of d'_D were smaller than those for d'_S for the both SIT models and the MIT, noise-limited model, reversing the expected pattern of discriminability for targets among similar and dissimilar distractors. Allowing d' s to scale with set size produced less variable λ estimates, and the expected relationship between d'_S , and d'_D for the MIT model only. AIC values were also smaller for the MIT than SIT models, with the smallest value obtained under the MIT, capacity-limited model – indicating a better fit between the observed and predicted hits and false

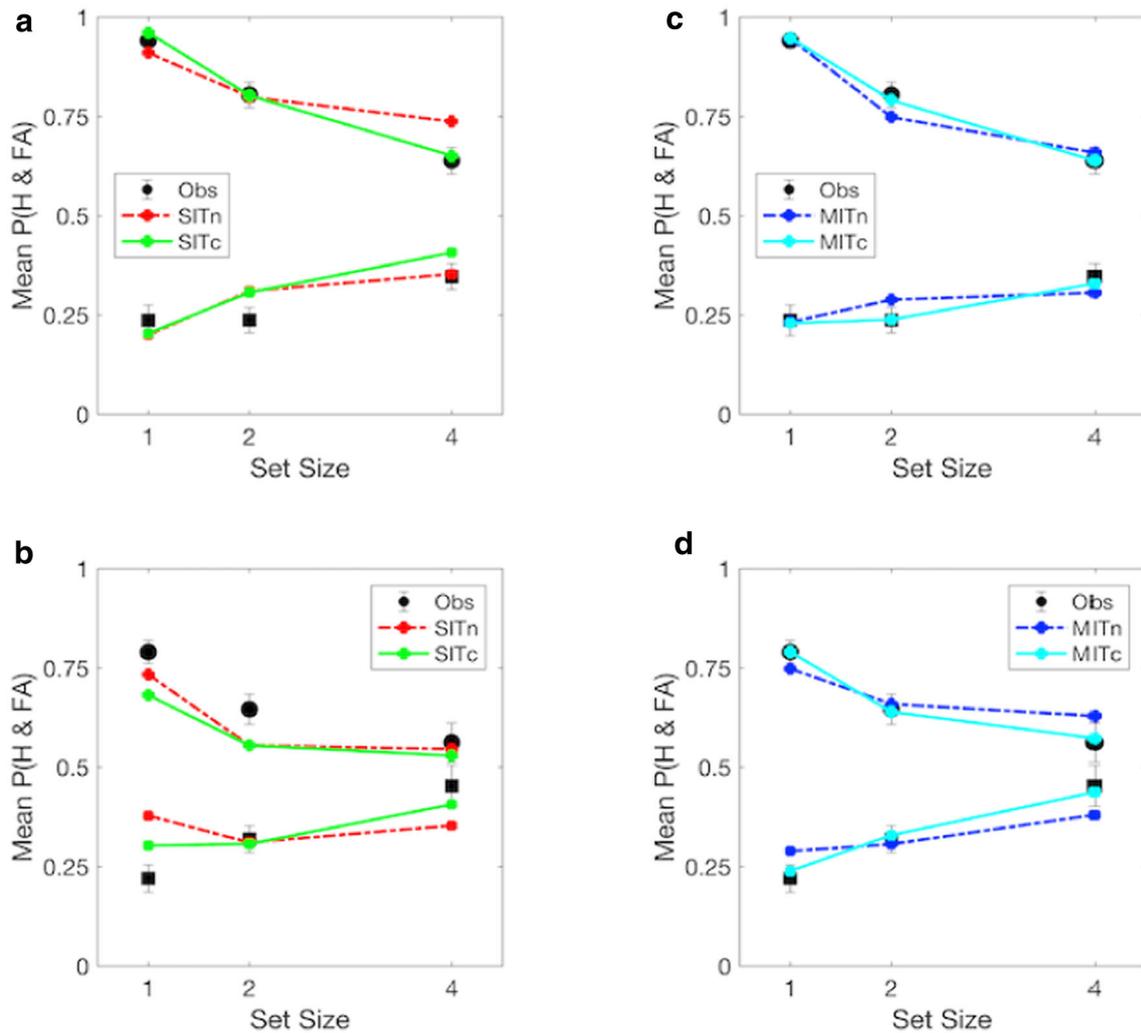


Fig. 4 Mean proportion of observed and estimated hits (circles) and false alarms (squares) by set size and search type. Single-item template, noise- (SITn) and capacity-limited (SITc) estimates on single- (a) and dual-target

(b) searches. Multiple-item template, noise- (MITn) and capacity-limited (MITc) estimates on single- (c) and dual-target searches (d). Error bars represent standard errors of observed means

alarms when accuracy was inversely scaled by the product of set size and the number of cued orientations. This group-level advantage was replicated at the individual level, where AIC values were smallest and goodness-of-fit indices largest for

the MIT, capacity-limited model for 10/11 observers. The mean absolute difference between observed and estimated data was also smaller for the MIT, capacity- (0.03) than the MIT, noise-limited (0.05), and SIT, capacity- (0.07) and noise-

Table 2 Mean (standard error) parameter estimates for the SIT and MIT noise- (n) and capacity-limited (c) models across equivalent set sizes in Experiment 1

	λ by search and number of comparisons				d'_S	d'_D	Fit	\sum AIC
	ST1	ST2-DT1	ST4-DT2	DT4				
SITn	-1.43 (.11)	-0.64 (.09)	-0.30 (.09)	-	-2.34 (.19)	-1.21 (.10)	< 0.001	1172.37
SITc	-2.05 (.19)	-1.56 (.17)	-1.36 (.18)	-	-2.94 (.25)	-2.33 (.17)	0.05 (.02)	1052.98
MITn	-1.94 (.25)	-.042 (.09)	-0.22 (.08)	0.22 (.11)	-2.72 (.29)	-1.05 (.11)	0.09 (.07)	932.84
MITc	-1.86 (.18)	-1.55 (.16)	-1.39 (.17)	-1.20 (.20)	-2.65 (.23)	-3.64 (.46)	0.38 (.10)	800.28

SIT models assume the number of comparisons is equal to display set size. On dual-target trials, DT4 = 8 comparisons. Subscripts S and D denote d' estimates for distractors with orientations similar and dissimilar to the target appearing in the display

AIC Akaike Information Criterion

limited (0.06) models (see Tables S1a to S1d in the ESM for individual estimates and fit-indices).

Discussion

The results above are consistent with previous demonstrations of the dual-target cost; observers were significantly less accurate on dual- compared to single-target searches across set sizes 1, 2 and 4. Our data extend previous findings from free-view (i.e., Barrett & Zobay, 2014; Menneer et al., 2007) to brief displays that preclude eye movements during single- and dual-target search. To determine whether the reduction in accuracy is consistent with comparisons based on one or two cued-orientations, we contrasted the observed probability distributions of hits and false alarms with those predicted by SIT and MIT noise- and capacity-limited SDT models of search. These revealed poor fits for the SIT models; indicating accuracy exceeding that predicted by a single-item limit on selection or decision criteria during search. In contrast, both MIT models provide a good approximation of observed accuracy distributions on single- and dual-target searches (see Fig. 4), with the range of estimated parameters and statistical comparisons favouring the capacity- over the noise-limited model at both individual and group-level analyses. The close fit between the observed and predicted data for the MIT capacity-limited model indicates the dual-target cost is (1) best approximated by an inverse relationship between accuracy and the product of set size and the number of cued orientations and (2) exceeds that predicted by an increase in decision-noise alone during independent searches for each cued-target. This result suggests the distinction between equivalent electrophysiological responses for independently cued targets (e.g. Grubert & Eimer, 2016) and behavioural indices of the dual-target cost (e.g. Irons et al., 2012) is attributable to increases in decision-noise as well as the distribution of capacity-limited decision resources when observers evaluate the presence of more than one potential target during search.

SDT models of single-target search have been successfully used to characterise the relationships between decision-noise, discriminability and set size for different stimuli (e.g. Eckstein et al., 2000; Pöder, 2017; Shaw, 1980; Smith, 2010; Verghese, 2001). Our analyses suggest extending set size to quantify the number of comparisons between items in VSTM (i.e., memory set size) and objects in the display is sufficient to explain the dual-target cost in brief displays. In information theoretic terms, capacity-limited models consider display duration a limit on the accumulation of evidence during search, which is divided by the number of objects in the display (sample-size model). Our data suggest the limit on accumulation can be described by scaling d' by the square root of the number of comparisons required to classify the display (i.e., $d'/\sqrt{(\text{attentional templates} * \text{set size})}$). This interpretation is

consistent with capacity-limited decision processes based on the concurrent evaluation of similarity distributions for independently cued targets.

Previous research suggests perceptual decisions and eye movements are both informed by an evaluation of the maximum sensory evidence of the target (Beutter, Eckstein, & Stone, 2003; Najemnik & Geisler, 2008). Other findings, however, have highlighted the parallel nature of saccadic programming and the potential for competition between neurons coding the locations of different stimuli during the planning and execution of eye movements (McPeck & Keller, 2001; Theeuwes, Kramer, Hahn, & Irwin, 1998). Saccadic latencies to cued-targets are slowed by the presentation of a distractor with cued features (Ludwig & Gilchrist, 2003), and presenting two cued targets during saccades elicits near-equivalent proportions of fixations to each target (Beck & Hollingworth, 2017). Shifts of attention during presaccadic planning are also thought to entail VSTM resources during the retention and evaluation of information used to evaluate postsaccadic visual input (Deubel & Schneider, 1996). The foveation of likely targets also requires the translation of similarity values from a parallel to a graded-serial representation during the planning and execution of saccades. Applied to dual-target search, these findings suggest competition between similarity maxima for independently cued features may disrupt processes that are specific to saccadic targeting during search. Experiment 2 was designed to establish whether the dual-target cost in saccadic targeting is consistent with that predicted by (1) a single-item limit on selection or (2) selection based on a multiple-item template during search for one of two potential targets.

Experiment 2

Method

Observers

Twelve observers were recruited to the study: Two were male and age ranged from 18 to 52 years ($M_{\text{age}} = 22.70$, $SD = 1.81$). All reported normal or normal-to-corrected visual acuity.

Apparatus

The experiment was run on an IBM PC with a 21-in. HP Trinitron P1130 monitor CRT monitor (Walnut, CA, USA). The display resolution was $1,280 \times 1,024$ pixels and the frame rate 100 Hz. Stimulus presentation and data collection were controlled using custom-built software in MATLAB (Mathworks, Natick, MA, USA) with Psychophysics toolbox extensions (Brainard & Vision, 1997; Kleiner et al., 2007). Viewing distance was maintained at 57 cm using a fixed chin rest and eye movements

were recorded with an EyeLink 1000 video-based eye tracker (SR Research Ltd., Ottawa, ON, Canada) at a sampling rate of 1,000 Hz and spatial resolution of $< 0.02^\circ$.

Stimuli

Displays contained four red ($x = 0.566, y = 0.333, Y = 5.49 \text{ cd/m}^2$) Landholt's Cs presented on a black background ($x = 0.294, y = 0.283, Y = 0.82 \text{ cd/m}^2$). Cs subtended $5.0^\circ \times 5.0^\circ$ at equally spaced locations on the circumference of a virtual circle with a radius of 7.0° . In contrast to Experiment 1, stimuli were designed to be large enough to discriminate leftward and rightward orientation from the central fixation and placed at more eccentric locations that required orthogonal saccadic vectors. As set size was fixed at four objects, the ellipses used to control crowding across different set sizes in Experiment 1 were not included in displays.

Procedure

The experiment used a factorial design to manipulate two independent variables: Search type (single- or dual-target) and stimulus type (target, similar and dissimilar distractor). On each block, Target-Left and Target-Right orientations were sampled from 90° and $270^\circ \pm 15^\circ$ to 30° , respectively. Distractor orientations were assigned as Target-Left and -Right $\pm 25^\circ$. Target and distractor orientations were fixed and the order of presentation for each target and search type were randomly assigned within each block.

Figure 5 illustrates the sequence of events on each trial. Trials began with a fixation cross at the centre of the screen. After 500 ms, the fixation was replaced by a cue containing two objects on the horizontal midline at $\pm 3.5^\circ$ from the centre

of the screen. On single-target searches, the cue consisted of one C-shape at the orientation assigned to Target-Left or Target-Right and one annulus. On dual-target searches the cue contained two C-shapes at the orientations assigned to Target-Left and Target-Right. Single- and dual-target cues were presented for 500 ms and 1,000 ms, respectively, before being replaced by a fixation cross for 1,000 ms. Search displays followed and always contained four objects – a cued target and one similar and two dissimilar distractors – which were randomly assigned to the four possible locations. Observers were instructed to fixate the object that matched a cue, and responses were measured using a box criterion to record the first fixation falling within the $5.0^\circ \times 5.0^\circ$ area surrounding each C-shape in the display. This method excludes saccades falling outside stimulus locations but produces results similar to the nearest endpoint criteria (Eckstein, Beutter, & Stone, 2001) and provides unambiguous indices of object selection during search. Once a response had been recorded, the bounding square surrounding the fixated object was coloured red for 50 ms and the search display was replaced by a blank screen for 500 ms.

Observers completed a short practice (20 trials) followed by four blocks of 60 trials in a single experimental session. Before each block, a 5-point calibration sequence was performed to ensure fixations to each C-shape could be recorded. Following successful calibration (gaze errors $< 0.5^\circ$), observers were prompted to press a key to begin each block of trials.

SDT models of saccadic targeting

The observer's task on each trial is to indicate which of four stimuli match the cue by fixating the target. Our statistical

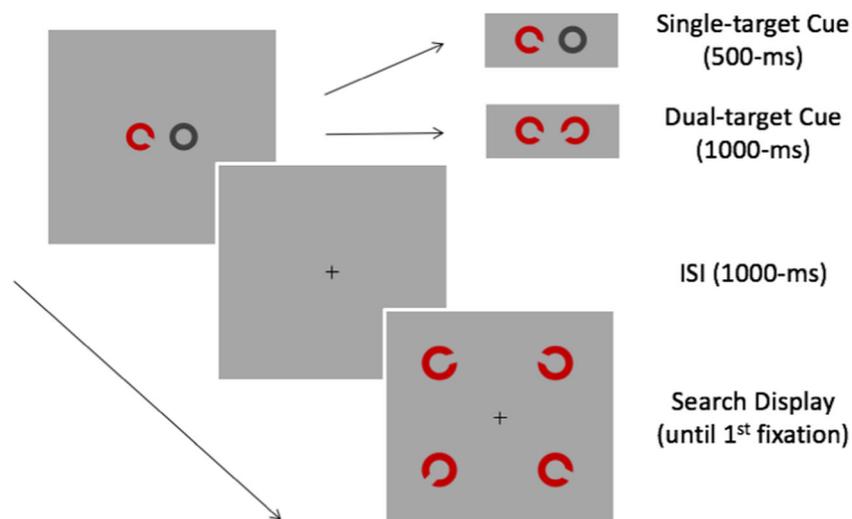


Fig. 5 Sequence of events on each trial of the overt search task. In this display, the target and similar distractor are in the bottom-right and top-left corners, respectively

description of this task uses the SDT framework derived for Experiment 1. Following the standard approach to modelling forced-choice tasks (Green & Swets, 1974; Kingdom & Prins, 2010; Wickens, 2002), we assume that the subject compares cued objects ($i = 1$ to 2) with those in the display ($j = 1$ to 4) and fixates the comparison that produces the maximum similarity s_{ij} to select the most likely target. The probability of the cue-object comparison (i, j) to yield maximum similarity is given by

$$P(s_{i,j} \text{ max}) = \int_{-\infty}^{\infty} p(y|C(i, j), D, T) \prod_{(i', j') \neq (i, j)} P(s_{i', j'} \leq y | C(i', j'), D, T) dy. \tag{4}$$

In Eq. 4, $p(y|C(i, j), D, T)$ denotes the probability density associated with $P(s_{ij} \leq y | C(i, j), D, T)$. The equation can be interpreted as follows: The product term in the integrand describes the probability that all comparisons different from (i, j) have results less than y . Weighing this value with the probability density for obtaining response y in the (i, j) comparison and integrating over all possible values of y yields the total probability for selecting comparison (i, j).

With D fixed to 4 and setting $\mu(C=t, D, T) = 0$ as in Experiment 1, the model has four free parameters, which represent the perceived similarity of similar (s) and dissimilar (d) distractors on single- (ST) and dual-target (DT) searches; $d'(s_{ST})$, $d'(d_{ST})$, $d'(s_{DT})$ and $d'(d_{DT})$. For each observer, there are four observed quantities, i.e., the probabilities for selecting the target and similar distractor on single- and dual-target searches, respectively (the probability for choosing the dissimilar distractor is fixed because of $P(t)+P(s)+2P(d) = 1$). Without any further restrictions on the d' parameters, the model is fully identified and would reproduce the observations perfectly. However, in analogy to Experiment 1, we consider noise-limited SIT and MIT models with $d'(s_{ST}) = d'(s_{DT}) = d'_S$ and $d'(d_{ST}) = d'(d_{DT}) = d'_D$. We also compute estimates for a capacity-limited MIT model in which d' scales with the number of comparisons required to search for one or two cued-targets (i.e., $d'(s_{ST}) = d'_S$, $d'(s_{DT}) = d'_S \sqrt{2}$ and $d'(d_{ST}) = d'(d_{DT}) = d'_D$). As set size is fixed across search types in this experiment, scaling d' s by $\sqrt{2}$ models a capacity limit that is specific to the MIT model when VSTM resources are distributed across independent attentional templates during search. All three models have two free parameters and 2 error degrees of freedom at the subject level.

The subject does not directly report the chosen comparison (i, j) but only the corresponding display object j . Therefore, the statistical model has to provide expressions for $P(j)$ rather than $P(i, j)$. This does not cause any complications for single-target searches, as there is only one cue $i=1$ and $P(j) = P(i=1, j)$. For dual-target searches in the MIT model, saccades to object j could be based on the comparison for either cue, i.e., ($i=1, j$) or ($i=2, j$) so that the

corresponding probabilities have to be added: $P(j) = P(i=1, j)+P(i=2, j)$. For the SIT model, the overall probability of fixating the target in a DT search is given by

$$P(Target) = \frac{1}{2} \left(\int_{-\infty}^{\infty} \varphi(y) \Phi(y-d'_S) \Phi^2(y-d'_D) dy + \int_{-\infty}^{\infty} \varphi(y-d'_D) \Phi^2(y-d'_S) \Phi(y-d'_D) dy \right). \tag{5}$$

In this equation, the first integral denotes the probability of target fixation when the observer activates the attentional template matching the target in the display. The second integral denotes the probability of target fixation when the observer activates the template matching the cued target that is not presented in the display. On these trials, the probability of fixating the target is low because the comparison between the attentional template and the target is unlikely to yield the maximum similarity (i.e., TL cue against TR target). For MIT models, the overall probability of selecting the target object in the noise-limited model is given by

$$P(Target) = \int_{-\infty}^{\infty} \varphi(y) \Phi^3(y-d'_S) \Phi^4(y-d'_D) dy + \int_{-\infty}^{\infty} \varphi(y-d'_D) \Phi(y) \Phi^3(y-d'_S) \Phi^3(y-d'_D) dy. \tag{6}$$

This equation is interpreted as follows. On a dual-target search, the participant performs one target comparison, three similar-distractor comparisons and four dissimilar-distractor comparisons. The first integral describes the probability that the subject selects the comparison between the target cue and target object. The second pertains to choosing the comparison between the cued target that is not present and the target object, which is considered a dissimilar-distractor comparison (i.e., TL cue against TR target). Models are again fitted using maximum-likelihood estimation. For the computation of integrals, we make use of the fact that φ is a normal distribution and apply Gauss-Hermite quadrature (Press, Teukolsky, Vetterling, & Flannery, 2007). Mean estimates and overall goodness-of-fit statistics are reported in the text. Individual estimates and fit-statistics are provided in the [ESM](#).

Results

Accuracy

Trials in which fixations to a display object were recorded in less than 100 ms or longer than 3 standard deviations from the mean of single- and dual-target searches, respectively, were excluded from further analyses ($M = 0.04$). Accuracy for Target-Left and Target-Right were not significantly different, $t(11) = 0.07$, $p > 0.95$, Cohen's $d = 0.019$, and the following analyses collapse responses across both targets. Table 3 lists the mean proportion of recorded first fixations and saccadic latencies to targets and distractors during single- and dual-

Table 3 Mean proportion and median saccadic latencies (milliseconds) of first fixations to targets and distractors on single- and dual-target searches in Experiment 2. Distractors are labelled similar (DS) and dissimilar (DD) relative to the orientation of the cued-target presented in the display

Object	Single target			Dual target		
	T	DS	DD	T	DS	DD
P (fixate)	.71 (.04)	.18 (.03)	.05 (.01)	.47 (.04)	.16 (.02)	.19 (.01)
Mean SL	671 (.15)	617 (.19)	548 (.18)	647 (.14)	610 (.21)	651 (.15)

Standard errors of means are reported in parentheses

The proportion shown for dissimilar distractors (DDs) is half the proportion of fixations made to both dissimilar distractors in the display

target searches. Data for single-target searches reveal highly accurate saccadic guidance, with observers demonstrating a clear bias to fixate targets over distractors. Relative to the target presented in the display, observers were also more likely to fixate similar than dissimilar distractors, illustrating similarity-based competition for selection during search. On dual-target searches, the distribution of fixations to targets and distractors was more equal. Planned comparisons yielded a significant decrease in target-fixations, $t(11) = 16.61$, $p < 0.001$, Cohen's $d = 4.80$, and increase in fixations to dissimilar distractors, $t(11) = 12.32$, $p < 0.001$, Cohen's $d = 3.56$, on dual- compared to single-target searches. The proportion of fixations to similar distractors for each type of search was not significantly different, $t(11) = 0.52$, $p > 0.05$, Cohen's $d = 0.15$.

Saccadic latencies

Our focus in this study was to investigate changes in the accuracy of saccadic targeting during single- and dual-target searches. In Experiment 1, short-duration displays constrained the accumulation of information equally for both types of search. In the saccadic tasks, however, observers may compensate for changes in the fidelity of guidance by sampling information longer before initiating saccades. In order to assess potential speed-accuracy trade-offs, we computed median saccadic latencies to targets and distractors for single- and dual-target searches for each observer (lower row, Table 3). On single-target searches, saccades to dissimilar distractors were faster than those to similar distractors and targets. On dual-target searches, mean saccadic latencies to dissimilar distractors and targets were almost identical. A repeated-measures ANOVA yielded a significant main effect of object type, $F(2,22) = 3.79$, $p = 0.038$, $\eta_p^2 = .26$, and a significant Search by Object type interaction, $F(2,22) = 8.44$, $p = 0.002$, $\eta_p^2 = .43$. The main effect of search did not reach statistical significance, $F(1,11) = 4.57$, $p > 0.05$, $\eta_p^2 = .29$. *Post hoc* tests revealed median saccadic latencies to dissimilar distractors were significantly faster on single- than dual-target searches ($M = 103\text{ms}$, $p < 0.001$), while differences to targets ($M = 24\text{ms}$, $p > 0.05$) and similar distractors ($M = 1\text{ms}$, $p > 0.05$) were not statistically significant.

Model estimates and fits

Figure 6 plots the mean observed and predicted proportions of fixations to targets and distractors on single- and dual-target searches. As can be seen, all three models produce reasonable fits to the data, with SIT and MIT models capturing the difference in the discriminability of targets against similar and dissimilar distractors on single-target searches, as well as the increase in the likelihood of fixating dissimilar distractors on dual-target searches. All three models also capture the marked decrease in the accuracy of first saccades on dual- compared to single-target searches. For single-target searches, the MIT, noise-limited model underestimates the proportion of fixations to targets and overestimates those to similar distractors. The same model also overestimates the likelihood of fixating the target on dual-target searches. Fits for the SIT, noise-limited and MIT, capacity-limited models are much closer to the observed data. For the former, mean estimates fall within 1 standard error of the observed means for all but dissimilar distractors on dual-target searches. For the latter, mean estimates fall within 1 standard error of observed means for all objects across both types of search. These differences are reflected in the statistical comparisons listed in Table 4, which yield smaller summed AIC values and higher fit indices for the MIT, capacity-limited than the SIT and MIT, noise-limited models. Fit indices also favour the MIT, capacity-limited model at the individual level, where AIC values were smaller and fit indices larger compared to the SIT and MIT, noise-limited models for 8/12 observers. The mean absolute difference between observed and estimated data was also smaller for the MIT-capacity (0.02) than the MIT, noise-limited (0.03) and SIT, noise-limited models (0.03; individual parameter estimates and goodness-of-fit statistics are presented in Tables S2a, S2b and S2c in the ESM).

Discussion

The results of Experiment 2 reveal a significant reduction in the accuracy of first saccades on dual- compared to single-target searches. This finding is consistent with previous evidence for a decrease in saccadic selectivity during dual-target

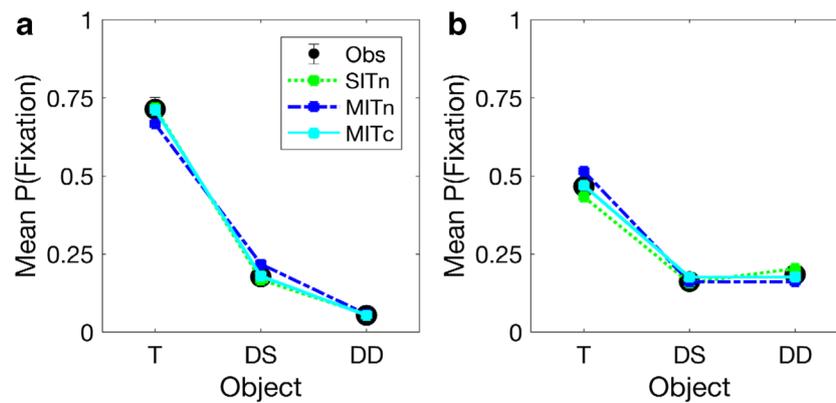


Fig. 6 Mean proportion of observed (Obs) and estimated fixations to targets (T), similar (DS) and dissimilar (DD) distractors on single-target (a) and dual-target (b) searches in Experiment 2. *SITn* single-item

template, noise-limited, *MITn* multiple-item template, noise-limited, *MITc* multiple-item template, capacity-limited model. Error bars denote standard errors of observed means

searches for colour-defined targets (Stroud, Menner, Cave, Donnelly, & Rayner, 2011; but see Beck et al., 2012). In the current study, changes in the proportions of fixations to objects classified in terms of their similarity to the target(s) provide an index of the localised selection that guides search. On single-target searches, 89% of fixations were made to objects that grouped with the cue; 71% to the target and 18% to the similar distractor. Saccades to each of the remaining dissimilar distractors averaged just 5%, indicating observers’ ability to select objects with similar over dissimilar orientations to the cue during search. On dual-target searches, 63% of saccades were made to objects that grouped with the cued-target that was presented in the display, 47% to the target and 16% to the similar distractor. Saccades to each of the dissimilar distractors increased to 19% of trials, a figure equivalent to that for similar distractors on single-target searches. In the current study, this equivalence is predicted by cue-based selection that reverses the similarity values assigned to leftward and rightward distractors during dual-target searches. Figure 1 illustrates this situation when Target-Left is cued and present in the display. Hypothetical similarity values for objects are plotted on the ordinate of panels B to D. Observers saccade to the location with the maximum value, with competition for selection reflecting the similarity between the cue and objects in the display. On a single target search (Panel B), competition is

largest between the cued target (TL) and its similar distractor (DL). On a dual-target search (Panel D), activation associated with the template for the rightward leaning target (TR) prioritises the complementary subset of distractors, which generates equivalent competition from similar (DL) and dissimilar distractors (DR) during saccadic targeting.

Figure 6 shows all three models provide reasonable fits to the observed distribution of saccades on single- and dual-target searches at the group level. Comparisons indicate better fit indices for the MIT, capacity-limited than the MIT, noise-limited model, where d' is independent of the number of comparisons required to localise the target. Group and individual comparisons also reveal better fit indices for the MIT, capacity-limited than the SIT, noise-limited model, supporting saccadic targeting based on separate similarity distributions for each cue. As with Experiment 1, comparisons of the fit indices for SIT and MIT models support a dual-target cost that is (1) best approximated by an inverse relationship between accuracy and the product of set size and the number of cued orientations, and (2) exceeds that predicted by an increase in decision-noise alone during search for independently cued targets. The analysis of saccadic latencies also rules out a general slowing or speed-accuracy trade off on dual- compared to single-target searches (Godwin, Walenckok, Houpt, Hout & Goldinger, 2015). Surprisingly, mean saccadic latencies to targets and distractors were slower on dual- than single-target searches, raising the possibility that observers require less evidence to initiate saccades when similarity distributions contain maxima for independently cued targets. Notably, however, the reduction for targets and similar distractors was not statistically significant, indicating any change in decision thresholds for these objects is relatively small. Despite the lack of reliable changes in latencies for these objects, our analyses did reveal a statistically significant increase in saccadic latencies to dissimilar distractors on dual- compared to single-target searches. This is likely to reflect first fixations to dissimilar distractors that were initiated before top-down selection was fully implemented

Table 4 Mean parameter estimates for the SIT, noise-limited (SITn), MIT, noise-limited (MITn) and MIT, capacity-limited- (MITc) models in Experiment 2. Standard errors of means are reported in parentheses and d' subscripts S and D denote similar and dissimilar distractors, respectively

Model	d'_S	d'_D	Fit	\sum AIC
SITn	-1.28 (.18)	-1.93 (.15)	.22 (.09)	382.96
MITn	-0.96 (.13)	-1.87 (.16)	.09 (.03)	392.78
MITc	-1.20 (.16)	-1.98 (.17)	.44 (.09)	331.22

SITn single-item template, noise-limited, *MITn* multiple-item template, noise-limited, *MITc* multiple-item template, capacity-limited model

during search, as well as an increase in the competitive weights associated with dissimilar distractors during dual-target search (van Zoest, Donk, & Theeuwes, 2004). Taken together, the accuracy and latency data are consistent with cue-related changes in the likelihood and latency of fixations to leftward and rightward leaning distractors when searching for two targets selects complementary sets of similarity distributions for objects in the display.

General discussion

The current study was designed to investigate the nature of the dual-target cost on perceptual decisions and saccadic targeting during covert and overt search. To do this, we applied SIT and MIT SDT models to predict observers' accuracy when comparisons between visual input and VSTM are limited to (1) a single-item attentional template or (2) multiple-item templates for independently cued targets during dual-target search. In addition, we compared observers' accuracy to noise- and capacity-limited derivations for both models of attentional control. Each assumes search entails the comparison of noisy internal representations of visual input with one or two cued-orientations. Noise-limited derivations model the dual-target cost as an increase in stochastic noise when the number of comparisons doubles on dual- compared to single-target searches. Capacity-limited derivations apply an inverse square-root relationship to model an additional dependency between discriminability and the number of comparisons during search (Corbett & Smith, 2017; Eckstein et al., 2000; Shaw, 1980; Palmer, 1994). Each provides a means to generalise EEG evidence for parallel top-down activation of independently cued features (Christie, Livingstone, & McDonald, 2015; Grubert & Eimer, 2015) to behavioural data by incorporating (1) decision-noise and (2) the distribution of resources during the evaluation of evidence for one of two potential targets during dual-target search.

Experiment 1 contrasted the accuracy of perceptual decisions in brief displays during single- and dual-target searches. Display durations of less than 100 ms were used to impose a common temporal constraint and preclude eye movements on both types of search. Comparisons between the observed and estimated probability distributions of hits and false alarms revealed better fits for the MIT than SIT models, and a better fit for the MIT, capacity- than MIT, noise-limited model. The close approximation between observed and estimated data for the former indicates the dual-target cost is consistent with an increase in decision-noise and the distribution of capacity-limited decision processes when observers compare visual input with independently cued target values. This interpretation supports previous

evidence for the parallel selection of independently cued features during dual-target search (Eimer & Grubert, 2014; Grubert & Eimer, 2015), and suggests the dual-target cost reflects a quantitative rather than a qualitative reduction in the accuracy of perceptual decisions (Barrett & Zobay, 2014; Hayward, Walenchok, Houpt, Hout & Goldinger, 2015; Irons et al., 2012). Modelling the dual-target cost as an inverse function of the number of comparisons required to classify the display provides a mechanistic account of the discrepancy between electrophysiological and behavioural markers of the dual-target cost (Christie et al., 2015; Grubert & Eimer, 2015) that is consistent with theoretical accounts of resource distribution during guided search (Barrett, Shimozaki, Jensen, & Zobay, 2016; Corbett et al., 2017; Eckstein et al., 2000; Eckstein, 2011; Palmer, 1994; Pöder, 2017).

Experiment 2 adapted the SDT models for Experiment 1 to a 4-AFC task in which the accuracy of first fixations was the dependent measure. The data revealed a significant decrease in the accuracy of initial fixations during dual- compared to single-target searches, extending findings from free-view tasks (e.g. Stroud et al., 2012). In our data, the dual-target cost reflects a marked change in the distribution of fixation probabilities to targets and distractors during single- and dual-target search (see Fig. 5). On single-target searches, observers made very few (~5%) saccades to dissimilar distractors. These erroneous fixations had shorter latencies than those to targets or similar distractors, illustrating the near exclusion of dissimilar distractors as inspection time increased. On dual-target searches, the requirement to evaluate evidence for both targets elicited a more equal distribution of first fixations to leftward and rightward distractors. Assuming first fixations provide an index of latent selection during search, changes in the distribution of fixations evidence cue-related changes in the weights assigned to subsets of distractors during single- and dual-target searches. Comparisons between the observed and predicted distribution of fixations indicated SIT and MIT models provide a good approximation of observers' performance on single- and dual-target searches. Comparisons between models indicated the MIT models produced the closest fit between observed and estimated data for 8/12 observers. Scaling d' by a factor of the square root of the product of cued orientations and set size produced better fits than fixing d' to a single value in 7/8 observers whose data were best approximated by MIT models. These results are consistent with a dual-target cost for most observers that is commensurate with an increase in decision-noise and the distribution of capacity-limited resources when saccadic targeting is informed by multiple-item attentional templates (Sewell et al., 2014). A comparison of saccadic latencies also revealed a significant decrease on

dual- compared to single-target searches, which was specific to distractors dissimilar to the target presented in the display. These data support cue-related changes in the selection-weights assigned to subsets of distractors on dual-target searches, and illustrate comparable saccadic latencies to objects based on their similarity to independently cued orientations.

The current results are difficult to reconcile with a fixed single-item limit on top-down selection during search. For most observers, our data support perceptual decisions and saccadic targeting that are informed by independently cued features during dual-target search. For a minority of observers, however, perceptual decisions and saccadic targeting were best approximated by SIT models that restrict the comparison of visual input to a single cued item during search (Olivers et al., 2011; Houtkamp & Roelfsema, 2006, 2009). This inter-subject variability suggests behavioural indices of the dual-target cost may reflect distinct post-selection strategies during the evaluation of evidence for one of two cued targets. In our models, d' s representing the similarity between cued orientations and objects in the display are fixed. Changes in accuracy predicted by each model therefore reflect changes in the way selected information is evaluated during single- and dual-target searches, rather than changes to localised responses that signal target features in the display (Cohen et al., 2009; Eimer, 1996; Luck & Hillyard, 1994). In terms of the theoretical distinction between guidance and target-identification during search (e.g. Castelano, Pollatsek, & Cave, 2008; Hout & Goldinger, 2015; Palmer, 1995), our analyses indicate the dual-target cost can be explained by a decrease in the accuracy of target-identification as well as changes in way selected information is compared to the attentional template during search. Determining the internal (i.e., VSTM capacity) and external (i.e., number of cued values, object complexity and target-distractor discriminability) factors that influence these strategies remains an important question for future research. A relevant distinction between this and previous studies of the dual-target cost is our use of targets that were matched in terms of perceptual similarity to numerically equivalent subsets of distractors. This design negates potential stimulus-driven asymmetries in the similarity distributions elicited by separate cues, which may have biased selection and evaluation in previous studies toward the most discriminable (or memorable) categorical target (i.e., Houtkamp & Roelfsema, 2009; Mestry, Menner, Cave, Godwin, & Donnelly, 2017).

Evidence that observers can evaluate evidence for independently cued targets in our study provides insights into the way decision-noise and capacity limits affect search accuracy in brief displays or during initial saccadic targeting. Our findings, however, are insensitive to dual-

target costs that may arise during subsequent saccadic sampling in free-view displays. Previous eye-movement data suggest observers can switch between simultaneous and successive modes of guidance during dual-target searches (Beck et al., 2012; Cave et al., 2018). In this situation, the reduction in the accuracy of saccades appears to reflect an increased number of fixations to objects dissimilar to both cues (Stroud et al., 2012). Whether this cost is commensurate with that evidenced in the current study, or reflects additional constraints associated with the withdrawal of resources from one attentional template during the planning of successive saccades (Rajsic et al., 2017), or the requirement to retain the locations and status of previously fixated objects against multiple decision criteria, remain important questions for future research.

Our results suggest an important determinant of behavioural measures of the dual-target cost is the requirement to distribute post-selection resources during the comparison of visual input with independently cued features in VSTM (Sewell et al., 2014; Smith & Sewell, 2013). In the current study, displays contained a maximum of four objects, a number that corresponds to estimates of VSTM capacity (see Fukuda, Awh, & Vogel, 2010). Notably, however, evaluating four objects against two target templates required eight comparisons, which exceeds VSTM estimates on most tasks. Recent evidence from hybrid search has shown observers can successfully detect one of 100 potential targets encoded in long-term memory (Drew, Boettcher, & Wolfe, 2017). Eye movements during hybrid search reveal a logarithmic relationship between dwell time and the number of potential targets, which is consistent with the comparison of visual input with multiple memory representations within a single fixation. Findings from hybrid search highlight potential limits on the timing and accuracy of successive fixations during search for categorical objects in long-term memory. Our results quantify the dual-target cost on the evaluation of independently cued feature values during perceptual decisions in the absence of eye movements and on the accuracy of first saccades during overt search. The ability of the SIT and MIT models to predict observers' performance across these tasks with relatively few free parameters suggests they provide a theoretically plausible description of attentional selection and the decision processes underlying the initial stages of dual-target search. As such, our results provide a bridge between electrophysiological markers of top-down activation for independently cued features and behavioural measures of the dual-target cost, and provide important insights into the relationship between objects in VSTM, attentional control and decision processes during covert search and initial saccadic targeting in free-view displays.

Data Availability Experiments reported in the manuscript were not preregistered. Summary data for Experiments 1 and 2 are available at: 10.25392/leicester.data.7955969.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Adamo, M., Wozny, S., Pratt, J., & Ferber, S. (2010). Parallel, independent attentional control settings for colours and shapes. *Attention, Perception, & Psychophysics*, *72*(7), 1730–1735.
- Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: Links, causes and implications for spatial attention. *Trends in Cognitive Sciences*, *10*(3), 124–130.
- Barrett, D. J., Shimozaki, S. S., Jensen, S., & Zobay, O. (2016). Visuospatial working memory mediates inhibitory and facilitatory guidance in preview search. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1533–1546.
- Barrett, D. J., & Zobay, O. (2014). Attentional control via parallel target-templates in dual-target search. *PLoS One*, *9*(1), e86848.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851–854.
- Beck, V. M., & Hollingworth, A. (2017). Competition in saccade target selection reveals attentional guidance by simultaneously active working memory representations. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(2), 225–230.
- Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science*, *23*(8), 887–898.
- Beutter, B. R., Eckstein, M. P., & Stone, L. S. (2003). Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *JOSA A*, *20*(7), 1341–1355.
- Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, *33*, 1–21.
- Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Bundesen, Claus, Thomas Habekost, and Søren Kyllingsbæk. A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review* *112.2* (2005): 291–328.
- Burmester, A., & Wallis, G. (2012). Contrasting predictions of low-and high-threshold models for the detection of changing visual features. *Perception*, *41*(5), 505–516.
- Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *Journal of Neuroscience*, *31*(25), 9315–9322.
- Castelhano, M. S., Pollatsek, A., & Cave, K. R. (2008). Typicality aids search for an unspecified target, but only in identification and not in attentional guidance. *Psychonomic Bulletin & Review*, *15*(4), 795–801.
- Cave, K. R., Menneer, T., Nomani, M. S., Stroud, M. J., & Donnelly, N. (2018). Dual target search is neither purely simultaneous nor purely successive. *Quarterly Journal of Experimental Psychology*, *71*(1), 169–178.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*(6), 2918–2940.
- Christie, G. J., Livingstone, A. C., & McDonald, J. J. (2015). Searching for inefficiency in visual search. *Journal of Cognitive Neuroscience*, *27*(1), 46–56.
- Cohen, J. Y., Heitz, R. P., Schall, J. D., & Woodman, G. F. (2009). On the origin of event-related potentials indexing covert attentional selection during visual search. *Journal of Neurophysiology*, *102*(4), 2375–2386.
- Corbett, E. A., & Smith, P. L. (2017). The magical number one-on-square-root-two: The double-target detection deficit in brief visual displays. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(7), 1376–1396.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, *248*(4962), 1556–1559.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837.
- Dombrowe, I., Donk, M., & Olivers, C. N. (2011). The costs of switching attentional sets. *Attention, Perception, & Psychophysics*, *73*(8), 2481–2488.
- Downing, P., & Dodds, C. (2004). Competition in visual working memory for control of search. *Visual Cognition*, *11*(6), 689–703.
- Drew, T., Boettcher, S. E., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye-tracking investigation of hybrid search. *Journal of Vision*, *17*(11), 5–5.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Eckstein, M. P. (1998). The Lower Visual Search Efficiency for Conjunctions Is Due to Noise and not Serial Attentional Processing. *Psychological Science* *9* (2):111–118
- Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11*(5), 14–14.
- Eckstein, M. P., Beutter, B. R., & Stone, L. S. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception*, *30*(11), 1389–1401.
- Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics*, *62*(3), 425–451.
- Eimer, M. (1996). ERP modulations indicate the selective processing of visual stimuli as a result of transient and sustained spatial attention. *Psychophysiology*, *33*(1), 13–21.
- Eimer, M. (2015). EPS Mid-Career Award 2014: The control of attention in visual search: Cognitive and neural mechanisms. *The Quarterly Journal of Experimental Psychology*, *68*(12), 2437–2463.
- Eimer, M., & Grubert, A. (2014). Spatial attention can be allocated rapidly and in parallel to new visual objects. *Current Biology*, *24*(2), 193–198.
- Ester, E. F., Sutterer, D. W., Serences, J. T., & Awh, E. (2016). Feature-selective attentional modulations in human frontoparietal cortex. *Journal of Neuroscience*, *36*(31), 8188–8199.
- Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 317.
- Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–182.
- Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1039–1054.
- Godwin, H.J., Walenchok, S.C., Houpt, J.W., Hout, M.C. and Goldinger, S.D. (2015) Faster than the speed of rejection: Object identification

- processes during visual search for multiple targets. *Journal of Experimental Psychology: Human Perception and Performance*, 41(4), 1007-1020.
- Green, D., & Swets, J. (1974). *Signal detection theory and psychophysics*. Melbourne, FL: Robert E. Kreiger Publishing Co., Inc.
- Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 86-101.
- Grubert, A., & Eimer, M. (2016). All set, indeed! N2pc components reveal simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, 42(8), 1215-1230.
- Hollingworth, A., & Beck, V. M. (2016). Memory-based attention capture when multiple items are maintained in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance* 42 (7):911-917
- Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*, 77(1), 128-149.
- Houtkamp, R., & Roelfsema, P. R. (2006). The effect of items in working memory on the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2), 423-442.
- Houtkamp, R., & Roelfsema, P. R. (2009). Matching of visual input to only one item at any one time. *Psychological Research*, 73(3), 317-326.
- Irons, J. L., Folk, C. L., & Remington, R. W. (2012). All set! Evidence of simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, 38(3), 758-775.
- Kaplan, I. T., & Carvellas, T. (1965). Scanning for multiple targets. *Perceptual and Motor Skills*, 21(1), 239-243.
- Kingdom, F. A., & Prins, N. (2010). *Psychophysics: A practical introduction*. London: Academic Press.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1-16.
- Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Research*, 49(10), 1194-1204.
- Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, 20(5), 1000-1014.
- Ludwig, C. J., & Gilchrist, I. D. (2003). Target similarity affects saccade curvature away from irrelevant onsets. *Experimental Brain Research*, 152(1), 60-69.
- 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., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, 29, 317-322.
- McPeck, R. M., & Keller, E. L. (2001). Short-term priming, concurrent processing, and saccade curvature during a target selection task in the monkey. *Vision Research*, 41(6), 785-800.
- Menner, T., Barrett, D. J., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology*, 21(7), 915-932.
- Menner, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: Effects of practice and target similarity. *Journal of Experimental Psychology: Applied*, 15(2), 125-139.
- Mestry, N., Menner, T., Cave, K. R., Godwin, H. J., & Donnelly, N. (2017). Dual-target cost in visual search for multiple unfamiliar faces. *Journal of Experimental Psychology: Human Perception and Performance*, 43(8), 1504-1519.
- Moore, K. S., & Weissman, D. H. (2011). Set-specific capture can be reduced by pre-emptively occupying a limited-capacity focus of attention. *Visual Cognition*, 19(4), 417-444.
- Moore, K. S., & Weissman, D. H. (2014). A bottleneck model of set-specific capture. *PLoS One*, 9(2), e88313.
- Müller, M. M., Andersen, S., Trujillo, N. J., Valdes-Sosa, P., Malinowski, P., & Hillyard, S. A. (2006). Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences*, 103(38), 14250-14254.
- Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, 8(3), 4.1-4.14.
- Olivers, C. N., & Eimer, M. (2011). On the difference between working memory and attentional set. *Neuropsychologia*, 49(6), 1553-1558.
- Olivers, C. N., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, 15(7), 327-334.
- Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, 34(13), 1703-1721.
- Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, 4(4), 118-123.
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40(10-12), 1227-1268.
- Pöder, E. (2017). Combining local and global limitations of visual search. *Journal of Vision*, 17(4), 10-10.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge, MA: Cambridge university press.
- Ptak, R. (2012). The frontoparietal attention network of the human brain: Action, saliency, and a priority map of the environment. *The Neuroscientist*, 18(5), 502-515.
- Rajsic, J., Ouslis, N. E., Wilson, D. E., & Pratt, J. (2017). Looking sharp: Becoming a search template boosts precision and stability in visual working memory. *Attention, Perception, & Psychophysics*, 79(6), 1643-1651.
- Roper, Z. J., & Vecera, S. P. (2012). Searching for two things at once: Establishment of multiple attentional control settings on a trial-by-trial basis. *Psychonomic Bulletin & Review*, 19(6), 1114-1121.
- Rutishauser, U., & Koch, C. (2007). Probabilistic modeling of eye movement data during conjunction search via feature-based attention. *Journal of Vision*, 7(6), 5-5.
- Salmela, V. R., & Saarinen, J. (2013). Detection of small orientation changes and the precision of visual working memory. *Vision Research*, 76, 17-24.
- Sewell, D. K., Lilburn, S. D., & Smith, P. L. (2014). An information capacity limitation of visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 40(6), 2214-2142.
- Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. *Attention and performance VIII*, 8, 277-295.
- Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology*, 14(3), 353-409.
- Smith, P. L. (2010). Spatial attention and the detection of weak visual signals. *Tutorials in visual cognition*, 211-259.
- Smith, P. L., Lilburn, S. D., Corbett, E. A., Sewell, D. K., & Kyllingsbæk, S. (2016). The attention-weighted sample-size model of visual short-term memory: Attention capture predicts resource allocation and memory load. *Cognitive Psychology*, 89, 71-105.

- Smith, P. L., & Sewell, D. K. (2013). A competitive interaction theory of attentional selection and decision making in brief, multielement displays. *Psychological Review*, *120*(3), 589–627.
- Stroud, M. J., Menneer, T., Cave, K. R., & Donnelly, N. (2012). Using the dual-target cost to explore the nature of search target representations. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 113–122.
- Stroud, M. J., Menneer, T., Cave, K. R., Donnelly, N., & Rayner, K. (2011). Search for multiple targets of different colours: Misguided eye movements reveal a reduction of colour selectivity. *Applied Cognitive Psychology*, *25*(6), 971–982.
- 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. *Psychological Science*, *9*(5), 379–385.
- van Moorselaar, D., Theeuwes, J., & Olivers, C. N. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1450–1464.
- 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.
- Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*(4), 523–535.
- Wickens, T. D. (2002). *Elementary signal detection theory*. New York: Oxford University Press.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11–11.
- Wing, A., & Allport, D. A. (1972). Multidimensional encoding of visual form. *Perception & Psychophysics*, *12*(6), 474–476.
- Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. *Integrated models of cognitive systems* (pp. 99–119).
- Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*, 154–64.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.