Mapping the timecourse of goal-directed attention to location and colour in human vision

Rachel C. Adams *, Christopher D. Chambers

School of Psychology, Cardiff University, United Kingdom

ARTICLE INFO

ABSTRACT

Goal-directed attention prioritises perception of task-relevant stimuli according to location, features, or onset time. In this study we compared the behavioural timecourse of goal-directed selection to locations and colours by varying the stimulus-onset asynchrony (SOA) between cue and target in a strategic cueing paradigm. Participants reported the presence or absence of a target following prior information regarding its location or colour. Results revealed that preparatory selection by colour is more effective at enhancing perceptual sensitivity than selection by location, even though both types of cue provided equivalent overall information. More detailed analysis revealed that this advantage arose due a limitation of spatial attention in maintaining a sufficiently broad focus ($>2°$) for target detection across multiple stimuli. In contrast, when target stimuli fell within 2° of the spatial attention spotlight, the strategic advantages and speed of spatial and colour attention were equated. Our findings are consistent with the conclusion that, under spatially optimal conditions, prior spatial and colour information are equally proficient at guiding top-down selection. When spatial locations are ambiguous, however, colour-based selection is the more efficient mechanism.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The amount of information in an everyday visual scene can be overwhelming. Due to our limited information processing capacity, attention is thus crucial for prioritising behaviourally relevant stimuli for further processing. A central issue in the psychology of goal-directed or 'top-down' selection is the relative primacy of different attentional systems. Here, in particular, we address the question of whether goal-directed attention is allocated primarily based on the location or colour of visual stimuli.

Popular theories of attention have employed a 'spotlight' (Posner, 1980; Posner, Snyder, & Davidson, 1980) or 'zoom lens' (Eriksen & St. James, 1986; Eriksen & Yeh, 1985) metaphor to describe the distribution of attention. According to this analogy, stimuli within the focus of attention receive enhanced visual processing, whereas stimuli falling outside of this region are processed to a lesser extent or suppressed. Gradient models further propose that the extent to which a stimulus is processed is inversely related to its distance from the attentional focus (LaBerge, 1983; Mangun & Hillyard, 1988; Shulman, Wilson, & Sheehy, 1985). These theories all emphasise the importance of location for visual selection.

Other theories, however, advocate the importance of features. The 'Dimensional Weighting' account of Müller et al. (2003) (see also Bundesen, 1990; Wolfe, 1994; and Wolfe, Butcher, Lee, & Hyle, 2003) proposes that top-down weights are assigned to stimulus dimensions according to their instructed relevance, thus allowing for a rapid search. Müller, Reimann, and Krummenacher (2003) showed that when participants are endogenously cued to a target colour or orientation, performance is faster on valid trials, and slower on invalid trials (compared to neutral trials), consistent with previous spatial cueing studies (Posner et al., 1980).

Although it is widely accepted that location improves visual selection (e.g. Cave & Pashler, 1995; Heinze et al., 1994; Hoffman & Nelson, 1981; LaBerge, 1983; Lamy & Tsal, 2001; Logan, 1996; Mangun & Hillyard, 1988; Posner, 1980; Posner et al., 1980; Shulman et al., 1985), the effects of feature-based attention appear less clear (Hillyard & Münte, 1984; Kasten & Navon, 2008; Theeuwes & Van der Burg, 2007; Vierck & Miller, 2008). Some studies have shown that the effects are generally smaller and less consistent (Juola, Bouwhuis, Cooper, & Warner, 1991; Laarni, 1999), while others have indicated that selection based on features occurs later, and is contingent upon selection by location (Anllo-Vento & Hillyard, 1996; Kim & Cave, 2001; Moore & Egeth, 1998; Shih & Sperling, 1996). The 'Feature Integration Theory' of Treisman and colleagues argues that spatial attention is a prerequisite for feature binding, and is therefore essential for object identification (e.g. Treisman & Gelade, 1980; Treisman & Gormican, 1988; Treisman & Sato, 1990).

* Corresponding author at: School of Psychology, Cardiff University, Park Place, Cardiff, CF10 3AT, United Kingdom. Tel.: +44 29 208 74007x77499.
E-mail address: adamsrc1@cardiff.ac.uk (R.C. Adams).

Other recent studies, however, have provided both electrophysiological and behavioural support for independent feature-based selection (Andersen, Müller, & Hillyard, 2009; Kasten & Navon, 2008; Saenz, Buracas, & Boynton, 2002). Treue and Martinez-Trujillo (1999) have not only demonstrated that motion processing can occur independently of location, but have also shown that spatial and feature-based attention are equally strong and additive within the receptive field. This observation gave rise to the 'Feature-Similarity Gain Model', in which the authors postulate that the gain of a sensory neuron reflects the similarity between that neuron's preferred feature and the behaviourally-relevant stimulus. Neuronal responses for the attended feature property are consequently increased, while those for other properties are decreased (Martinez-Trujillo & Treue, 2004). Evidence from functional magnetic resonance imaging (fMRI) studies illustrates that this mechanism of neuronal enhancement operates across the entire visual field (Saenz et al., 2002; Serences & Boynton, 2007).

Most researchers now acknowledge that both spatial and feature-based attention play a role in visual selection, yet it remains unclear which of these is the faster and more efficient mechanism. Studies utilising event-related potentials (ERPs) and visually evoked potentials (VEPs) have typically found that selection by location precedes selection based on features (Anllo-Vento & Hillyard, 1996; Harter, Aine, & Schroeder, 1982). These results suggest that spatial attention acts as an early mechanism to control neuronal gain, whereas feature-based processing takes longer to engage (see also Juola et al., 1991). Anllo-Vento and Hillyard (1996) further argued that feature-based selection requires feedback signals and occurs only once stimuli have already been selected based on their location.

In contrast, however, Zhang and Luck (2009) demonstrated that when attention is studied under conditions of heightened competition (with stimuli presented simultaneously rather than sequentially), colour-based attention can influence feed-forward visual processing within 100 ms of stimulus onset, even at unattended locations. Other recent evidence indicates that activity corresponding to feature-selection can precede that for location-selection (Hopf, Boelmans, Schoenfeld, Luck, & Heinze, 2004; Zanto, Rubens, Bollinger, & Gazzaley, 2010). Contrary to previous conclusions, this result would imply that visual selection is originally driven by task-relevant features, which are then used to guide spatial attention.

Using a purely behavioural paradigm, Lui, Stevens, and Carrasco (2007) explored the temporal dynamics of goal-directed spatial and feature-based (motion) attention by measuring their perceptual advantages across varying stimulus-onset asynchronies (SOAs) between an arrow cue and target. Importantly, their findings suggest that the allocation of spatial attention (150–300 ms) precedes that of feature-based attention (300–500 ms). The authors argued that this finding is consistent with the retinotopic organisation of the early visual cortex, with spatial information being more readily accessible and easily encoded than information based on motion direction. Information regarding specific features, on the other hand, must be retrieved from higher cortical areas and is likely to take longer to develop. Conversely, Schenkluhn, Ruff, Heinen, and Chambers (2008) found that at an SOA of 600 ms the attentional advantage conferred by colour cues was substantially greater than that for spatial cues. These findings accord with those of Giesbrecht, Weissman, Woldorff, and Mangun (2006), who also demonstrated that colour cueing produces more accurate performance than spatial cueing at longer SOAs (1–8 s).

Such observations, however, may be feature-specific. For example, previous studies have shown that colour cueing can lead to faster (Found & Müller, 1996; Müller et al., 2003) and more accurate (Theeuwes & Van der Burg, 2007) detection than orientation or shape cueing, respectively. Some researchers have also argued that not all features are capable of directing attention (Wolfe & Horowitz, 2004), but rather that feature-based attention could be limited to certain attributes that most commonly capture attention (Maunsell & Treue, 2006). Harter et al. (1982) argued that the relative processing time of various features

depends on their complexity. However, the effects of colour and motion-direction appear to be of similar magnitude (Liu, Hospadaruk, Zhu, & Gardner, 2011; Saenz et al., 2002; Wolfe & Horowitz, 2004).

Taken together these results suggest at least three potential hypotheses to explain differences in the efficacy of goal-directed spatial and colour-based attention. First, spatial attention may boost an earlier perceptual advantage while colour-based attention dominates at longer preparatory intervals (Fig. 1a). Alternatively, colour-based and spatial attention may differ in the rate at which attention benefits perception (slope), with information about the target colour providing a universally greater attentional benefit than information about the target location (Fig. 1b). Third, the overall perceptual gains of spatial and colour-based attention may be equivalent, but colour-based attention may simply accumulate faster (Fig. 1c).

The purpose of the current study was to contrast these hypotheses by probing the relative timecourse of spatial and colour-based selection using a strategic cueing paradigm. To ensure that selection was strictly goal-directed, we used arbitrary letters as colour and location cues because existing evidence suggests that the use of arrow cues (e.g. Ristic & Kingstone, 2006) or colour patches (Theeuwes & Van der Burg, 2007) can induce involuntary shifts of attention.

2. Method

2.1. Participants

Twelve healthy observers participated in the experiment (5 females). All had normal or corrected-to-normal vision and reported normal colour vision (aged 20–33 years, $M = 25$, $SD = 3.8$). All observers provided written consent prior to taking part and were debriefed and paid for their participation.

2.2. Apparatus

Participants were seated in a darkened testing chamber with their head fixed using a chin and forehead rest with temple stabilisers. Stimuli were presented via a 21 in. CRT monitor at a fixed distance of 65 cm (60 Hz vertical refresh rate, 1024×768 resolution). The experiment was programmed in Visual Basic 6.0. Eyetracking was undertaken with a 250 Hz Cambridge Research Systems Video Eyetracker Toolbox.

2.3. Stimuli and procedure

Fig. 2 shows the experimental display sequence. Each testing block was split according to attention condition to create two 'half-blocks'; this meant that participants undertook either a spatial half-block (intermingled with neutral trials) followed by a feature half-block (intermingled with neutral trials) or vice versa. The order of alternation was counterbalanced across participants. At the beginning of each half-block an instruction screen was presented with the cues and corresponding locations or colours. Observers were permitted unlimited time to learn the associations presented in the instruction screen; the following trial was self-initiated with a button press. Due to the arbitrary mapping of the cues to specific locations or colours, and to minimise working memory load, the instruction screen was repeated every 5 trials.

We aimed to ensure that attention was allocated purely strategically by using arbitrary letter cues. Eight consonants (D, F, Z, J and H, M, T, K) were mapped on to four colours (red, yellow, pink and blue) and four quadrants (upper left, upper right, lower left, lower right) so that each cue validly eliminated 75% of the distractors as possible targets. The letter mapping for either colours or locations was reversed for half of the observers (i.e. D, F, Z, J corresponded to colours for half the observers and locations for the remaining half); this ensured that, on average, the cue stimuli were identical in all conditions. The cue letters were selected