Color-selective attention need not be mediated by spatial attention

Søren K. Andersen
Institut für Psychologie I, Universität Leipzig, Leipzig, Germany

Matthias M. Müller
Institut für Psychologie I, Universität Leipzig, Leipzig, Germany

Steven A. Hillyard
Department of Neurosciences, University of California at San Diego, La Jolla, CA, USA

It is well-established that attention can select stimuli for preferential processing on the basis of non-spatial features such as color, orientation, or direction of motion. Evidence is mixed, however, as to whether feature-selective attention acts by increasing the signal strength of to-be-attended features irrespective of their spatial locations or whether it acts by guiding the spotlight of spatial attention to locations containing the relevant feature. To address this question, we designed a task in which feature-selective attention could not be mediated by spatial selection. Participants observed a display of intermingled dots of two colors, which rapidly and unpredictably changed positions, with the task of detecting brief intervals of reduced luminance of 20% of the dots of one or the other color. Both behavioral indices and electrophysiological measures of steady-state visual evoked potentials showed selectively enhanced processing of the attended-color items. The results demonstrate that feature-selective attention produces a sensory gain enhancement at early levels of the visual cortex that occurs without mediation by spatial attention.

Keywords: visual attention, steady-state visual evoked potential, feature selection, color


Introduction

Everyday experience informs us that color is a highly effective cue for locating relevant stimuli in a visual scene. Experiments using visual search tasks have verified that attention can be directed selectively to target items of a designated color, which can then be identified more rapidly and accurately than distractor items (e.g., Brawn & Snowden, 1999; Moore & Egeth, 1998). A controversy has arisen, however, regarding the mechanism by which color cues guide attention. On one hand, several lines of evidence have led to the proposal that color cues are effective because they allow spatial attention to be focused rapidly on their location; in other words, attention to color is primarily, if not exclusively, mediated by attention to location (Moore & Egeth, 1998; Pöder, 2001; Shih & Sterling, 1996; Tsal & Lavie, 1993). On the other hand, recent studies by Vierck and Miller (2005, 2008) have found that color-cued target items presented in a rapid serial visual presentation (RSVP) sequence with distractors at a common location were discriminated better than the uncued distractors. They concluded that stimuli may be directly selected for further perceptual processing without mediation by spatial selective attention.

Studies in which event-related brain potentials (ERPs) were recorded to items of relevant and irrelevant colors presented in a randomized RSVP sequence have provided evidence consistent with such a “direct selection by color” mechanism (Anllo-Vento & Hillyard, 1996; Anllo-Vento, Luck, & Hillyard, 1998; Hillyard & Münte, 1984; Karayanidis & Michie, 1996; Lange, Wijers, Mulder, & Mulder, 1998; Ruijter, De Ruiter, & Snel, 2000; Smid & Heinze, 1997; van der Stelt, Kok, Smulders, Snel, & Boudewijn Gunning, 1998; Vierck & Miller, 2008). In these studies, the relevant-color items elicited larger cortical ERPs in the latency range of 150–300 ms after stimulus onset than the irrelevant-color items presented at the same location. While these ERP experiments appear to demonstrate direct stimulus selection on the basis of color, it seems conceivable that attention to location might still play a role. In particular, the appearance of a relevant-color item in an RSVP sequence might serve as a cue to briefly intensify the focus of spatial attention either upon the item itself or upon its representation stored in visual short-term memory (see Shapiro & Terry, 1998 for a similar proposal in the context of attentional blink experiments). This proposal—that even “direct” color selection in an RSVP sequence as reported by Vierck and Miller (2005) may have a spatial basis—accords with findings that ERP
components elicited by relevant-color stimuli were spatio-
topically lateralized to the contralateral hemisphere when stimulus sequences were presented at an eccentric location (Anllo-Vento & Hillyard, 1996; Hillyard & Münte, 1984; Karayanidis & Michie, 1996).

Recent neuroimaging and electrophysiological studies have provided additional evidence that selection of a relevant color may be dissociated from spatial attention. In an fMRI experiment, Saenz, Buracas, and Boynton (2002) presented an array of intermingled red and green dots to one visual field, with either the red or the green dots designated as relevant, while an array of dots that were either all green or all red was presented to the opposite visual field. It was found that the contralateral dot arrays of the attended color elicited larger neural responses in multiple visual cortical areas than did dots of the irrelevant color, thus supporting a mechanism of global feature gain increase throughout the entire visual field (Maunsell & Treue, 2006; Treue & Martinez-Trujillo, 1999). Using a very similar design but recording ERPs, Zhang and Luck (2009) showed that flashed dot arrays of the attended color in the unattended visual field elicited a larger early (ca. 100 ms) neural response in the visual cortex than arrays of the unattended color; the authors concluded that color-based attention influenced the early feed-forward flow of attended-color information throughout the visual field. Neither of these studies, however, provided a direct demonstration that neural responses to items of the attended color within the attended array were differentially enhanced relative to the intermingled irrelevant-color items.

Evidence for enhanced color-selective processing within an attended array was obtained in studies where steady-state visual evoked potentials (SSVEPs) were recorded to large arrays of intermingled red and blue items that were flashed at different rates (e.g., red at 7.0 Hz and blue at 11.7 Hz; Andersen, Hillyard, & Müller, 2008; Müller et al., 2006). When attention was directed to items of one color, the amplitude of the frequency-tagged oscillatory SSVEP elicited by that color in early visual cortical areas was enlarged relative to when the other color was attended. The authors proposed that this amplification of neural activity elicited by the attended items reflects the operation of a color selection mechanism that can be applied across a visual field cluttered with intermingled attended items and distractors. Such a feature enhancement mechanism could enable the rapid identification of feature conjunction stimuli during visual search as proposed by “guided search” models (Wolfe, 1998).

In the above-described SSVEP experiments (Andersen et al., 2008; Müller et al., 2006), each of the red and blue items (125–150 of each) was continuously shifted in position by a small distance in a random walk as they flickered in order to discourage spatial tracking of individual items. Even though these random shifts of position occurred rapidly, at either 30 Hz (Müller et al., 2006) or 120 Hz (Andersen et al., 2008), it is still conceivable that subjects could have adopted a “multiple spotlight” strategy and tracked the rapid, jerky movements of some subset of the relevant-color items. The present SSVEP experiment was designed to rule out any possibility that spatial attention could be selectively allocated to the locations of the relevant- versus irrelevant-color items. This was achieved by shifting the position of the intermingled red and blue dots (100 of each) to an entirely new random location after each flicker cycle (at 10–12 Hz) rather than shifting each dot by a small amount on each cycle, which might have allowed tracking. In this way, attention could be directed to either the red or the blue dot populations while precluding selection on the basis of location or on the basis of focusing attention in time.

### Methods

#### Participants

Seventeen subjects (13 females, 1 left-handed, age range 19 to 28 years, mean 23.2 years) with normal color vision and normal or corrected-to-normal visual acuity participated after giving informed consent. Two participants were subsequently excluded from the sample due to inadequate task performance in the staircase procedure performed prior to EEG recordings (see below).

#### Stimuli

Two overlapping and interdigitated random dot kinematograms (RDKs) of different colors (red and blue) were presented on a gray background (Figure 1). Each RDK flickered with a 50/50 on/off cycle at a specific frequency (red: 10 Hz, blue: 12 Hz), thereby eliciting distinctive, frequency-tagged SSVEPs. All stimulations were synchronized to the screen refresh rate of 120 Hz. The RDKs had a diameter corresponding to 12.78° of visual angle, and each consisted of 100 randomly distributed dots, each subtending 0.32° × 0.32°. Dots unpredictably changed their positions in synchrony with the flicker (10 or 12 Hz), i.e., for each cycle of the flicker a new random sample of positions was drawn for all dots of the respective RDK. Dots never overlapped with one another so as to prevent any possible induction of a depth cue. Prior to recordings, the luminance of the red and blue dots was equated for each subject to the luminance of the gray background (9.2 cd/m²) by means of heterochromatic flicker photometry. Stimuli were presented on a 19-inch cathode ray tube monitor at a resolution of 640 × 480 pixels and 32 bits per pixel color resolution viewed at a distance of 80 cm.

Each trial started with the presentation of a color cue in the shape of a fixation cross for a randomly chosen interval of 600 to 900 ms. The cue could have one of three colors, indicating in which RDK the targets were most
likely to appear in the following trial: a red or blue cue indicated the RDK of the respective color, while a magenta cue indicated an equal probability of targets appearing in either the red or blue RDK. After the cue interval, the fixation cross turned gray, and the red and blue flickering dots appeared and remained on screen for 3042 ms. Before the start of the succeeding trial, a gray background was presented for 900 ms.

Participants were given the task of detecting brief intervals (200 ms) of reduced luminance of 20% of the dots (targets), which could occur unpredictably in either RDK. Target detections were reported by pressing a button, with equal emphasis on speed and accuracy. Up to three targets could appear within a single trial. The onset of targets never occurred earlier than 350 ms post-stimulation-onset, and subsequent targets within a trial were always separated by at least 850 ms. In trials where red or blue was cued, 83% of targets appeared in the cued color (validly cued targets), while the other 17% of targets appeared in the uncued color (invalidly cued targets). The magenta (neutral) cue indicated a 50/50 chance of targets appearing in either the red or blue RDK.

Procedure

Participants received training on the target detection task for up to three blocks of 72 trials each prior to data recording. This was followed by three blocks of a staircase procedure, which adjusted task difficulty to an intermediate and equal level for red and blue targets. In the staircase procedure, participants performed the same task as during the training and the later data recording. The size of the luminance decrement of targets was adjusted separately for red and blue targets by interleaved two-up one-down staircases (Levitt, 1971) driven by detection responses to neutrally cued targets only. After the staircase procedure, six experimental blocks were administered. Training, staircase, and experimental blocks consisted of 72 trials each, sorted into 6 runs of 12 trials each under the same cue condition. Thus, within each block there were 2 runs of each cue condition, presented in randomized order. Runs of the same cue condition never followed each other directly. Fifty percent of all trials contained between one and three targets, while the rest contained none.

The experimental blocks included a total of 432 trials (6 blocks × 72 trials/block). The total number of targets presented was 144 for each cue condition (red cued: 120 red and 24 blue; blue cued: 24 red and 120 blue; both cued: 72 red and 72 blue).

Behavioral data

Detection responses occurring from 200 to 1000 ms post-target onset were counted as hits. Hit rates and reaction times were collapsed over red and blue targets and subjected to a repeated-measures ANOVA with the factor of cue type (valid, neutral, invalid). All responses occurring outside the above-mentioned response window following target onsets were counted as false alarms. The total number of false alarms was subjected to a repeated-measures ANOVA with the factor of cue type (red, blue, or neutral).

SSVEP recordings and analysis

During EEG recordings, participants were seated in a comfortable chair in a dimly lit electrically shielded chamber. Brain electrical activity was recorded from 64 Ag/AgCl electrodes mounted in an elastic cap using a BioSemi ActiveTwo amplifier system (BioSemi, Amsterdam, the Netherlands) set to a sampling rate of 256 Hz. Lateral eye movements were monitored with a bipolar outer canthus montage (horizontal electrooculogram). Vertical eye movements and blinks were monitored with a bipolar montage positioned below and above the right eye (vertical electrooculogram).

Processing of EEG data was performed using the EEGLab toolbox (Delorme & Makeig, 2004) in combination with custom routines written in Matlab (The Mathworks, Natick, MA). Analysis epochs began 100 ms before the onset of the flickering RDKs and extended for 3100 ms thereafter. Only trials without targets or distractors were included in the SSVEP analysis to ensure that attentional selection was based solely on color with no influence of luminance changes or target detection responses. Trials with eye movements or blinks were rejected from further analysis, and all remaining artifacts were corrected or rejected by means of an automatic procedure (SCADS, Junghöfer, Elbert, Tucker, & Rockstroh, 2000). After subsequent algebraic transformation of the EEG data to
average reference, all epochs within the same cue condition were averaged for each participant.

The SSVEP amplitude at each electrode was calculated by Fourier transformation of a time window from 400 to 2900 ms after stimulus onset. This time window was chosen in order to exclude the visual evoked potential to stimulus onset. Prior to Fourier transformation, data in this time window were detrended (removal of mean and linear trends) in order to correct for any linear drifts. SSVEP amplitudes were quantified as the absolute value of the complex Fourier coefficients at the two stimulation frequencies. Based on the topographical distribution of SSVEP amplitudes averaged over the three cue conditions, a cluster of 4 electrode sites (Oz, O1, O2, Iz) where amplitudes were maximal was chosen and amplitudes were averaged across these four electrodes for statistical analysis. The data processing steps described above yielded SSVEP amplitude values for each stimulation frequency (i.e., for each RDK) for each of the three cue conditions. The SSVEP amplitudes for each stimulation frequency were averaged across these four electrodes for statistical analysis.

Results

Behavioral data

Cueing condition (valid, neutral, invalid) influenced both hit rates ($F(2,28) = 7.69, p < 0.05, \eta^2 = 35.5\%$) and reaction times ($F(2,28) = 20.67, p < 0.0005, \eta^2 = 59.6\%$). Hit rates to invalidly cued targets were lower than hit rates to validly or neutrally cued targets (both $p < 0.05$). Reaction times were fastest when targets were validly cued, intermediate when they were neutrally cued, and slowest when they were invalidly cued (all contrasts: $p < 0.005$). The number of false alarms was generally very low (averages: red cued: 2.13, blue cued: 2.53, both cued: 2.13) and did not differ between conditions ($F(2,28) = 0.17, p > 0.1$).

SSVEP amplitudes

Figure 3A depicts the grand-average SSVEP amplitude spectra for the three attention conditions averaged across a cluster of 4 occipital electrodes. SSVEP amplitudes were highest when the driving RDK was attended, lowest when the other RDK was attended, and intermediate when both RDKs were attended. The statistical analysis of the normalized SSVEP amplitudes collapsed over both frequencies confirmed this pattern: Color-selective attention had a strong influence on SSVEP amplitudes ($F(2,28) = 25.54, p < 10^{-5}, \eta^2 = 64.6\%;$ all contrasts, $p < 0.005$).

To gain information about the cortical regions where sensory signals are amplified by feature-selective attention, cortical sources were localized by means of variable resolution electromagnetic tomography (VARETA, Bosch-Bayard et al., 2001) and statistically compared between the attend red and attend blue conditions. As depicted in Figure 3C, the cortical currents giving rise to the SSVEP attention effect (attended minus unattended) were localized to the posterior, medial occipital cortex. The MNI (Evans et al., 1993) coordinates of the maximum modulations with attention for the blue dots were −14 91 −10 (left) and 14 −91 −10 (right) and for the red dots −43 −84 −10 (left) and 14 −91 −10 (right). Thus, similar to our previous findings (Andersen et al., 2008; Müller et al., 2006) maximal modulation with attention was found in regions containing early visual areas V1, V2, and V3.

Discussion

As reviewed in the Introduction section, previous results from both behavioral and physiological studies are consistent with the hypothesis that stimuli can be selected directly on the basis of color cues without mediation by spatial attention. We would submit, however, that a definitive experimental test of this hypothesis must fulfill several stringent requirements, and that none of the previous studies was fully compliant. First, the stimuli should be arranged so that items of relevant and irrelevant colors are not spatially separated, a point emphasized by
Vierck and Miller (2005, 2007). Second, the items of relevant and irrelevant colors should not be presented at separate times, which might allow for the focusing of spatial attention in time. Third, either behavioral or physiological evidence for selective processing of the attended-color items themselves must be obtained, with the strongest evidence being the joint presentation of both types of evidence obtained concurrently. To the best of our knowledge, the present study is the only one to meet all of these design requirements and thus provides the most definitive evidence to date that stimuli may be selected directly by virtue of their color without mediation by the focusing of spatial attention.

Behavioral and electrophysiological results from previous studies allow us to rule out the possibility that subjects selectively attended to the red and blue dots on the basis of flicker frequency rather than color. In a behavioral control experiment, we compared conditions in which dots flickered in the same frequency and phase or in different frequencies and found that flicker frequency did not provide a useful cue for attentional selection, even when the frequency difference was greater than in the present study (7.0 Hz vs. 11.7 Hz; Müller et al., 2006). Furthermore, the pattern of false alarms and the attentional modulations of SSVEP amplitudes observed in Andersen et al. (2008) were determined by the specific features that were attended and extended across stimuli presented at different frequencies. This pattern of results can be explained by selection of relevant features but is inconsistent with a selection based on flicker frequencies.

While the present results show that color selection can be direct, there is still much to be learned about the mechanism by which it is achieved. Vierck and Miller (2007) found no evidence that advance cueing of the most likely color of a single item produced any validity effects on discrimination speed or accuracy, which they took as evidence against a “signal enhancement” (faster and more accurate processing of the expected color item) mechanism for attention to color. Instead, they proposed that the effects of color-selective attention found in previous studies were likely to have resulted from a mechanism of uncertainty reduction at a decision stage, whereby more weight was given to the attended information. As Vierck and Miller pointed out, however, the assigned discrimination task in their experiments was independent of the color information, so their participants may simply have ignored the color cueing/expectancy manipulation. Indeed, we consider the enlarged SSVEP elicited by the attended-color items in early visual cortex in the present study to provide strong evidence for a signal enhancement mechanism. While the SSVEP lacks specific timing information, the attention-related augmentation of the evoked neural signals in the visual cortex is a clear index of enhanced early sensory processing of the attended-color items.

The present finding that color-selective attention results in enhanced processing of the selected items in the visual cortex is consistent with previous fMRI (Saenz et al., 2002).
and SSVEP (Andersen et al., 2008; Müller et al., 2006) studies showing that attended-color stimuli elicited stronger neural responses at an early stage of visual-cortical processing. As noted in the Introduction section, however, the designs of all these studies had some spatial separation between the intercalated attended and unattended items making it difficult to completely rule out spatial mediation of the color-selective processing. Nonetheless, the results of these previous studies and those of the present experiment may be readily interpreted in the framework of the feature-similarity gain model (Maunsell & Treue, 2006; Treue & Martinez-Trujillo, 1999), which specifies that all neurons that respond selectively to a particular feature (such as a particular color) will show an increase in gain or responsiveness when that feature is attended. This increase in gain may extend even to neurons with spatial receptive fields that are distant from the attended region (e.g., in the opposite visual field). Moreover, when the attended stimuli are defined as having two features (e.g., a particular color and shape), the sensory gain of cortical neurons is facilitated in a parallel, additive manner for each feature (Andersen et al., 2008; Bichot, Rossi, & Desimone, 2005). The neural circuitry that underlies such widespread facilitation of color processing has yet to be worked out, but presumably would involve top-down projections from attentional control areas that produce a differential gain change for neurons preferring the attended feature across the entire retinotopic representation (Maunsell & Treue, 2006). The present study affirms that such feature-selective gain enhancement can occur within an attended stimulus array and without mediation by spatial attention.

While the present results show that feature selection can occur in the absence of focused spatial attention on the relevant items, a more general role for spatial attention cannot be ruled out. It is possible, for example, that the color-selective processing demonstrated in the present study can only take place within a region of the visual field where spatial attention is diffusely distributed. Indeed, many of the participants in the present study reported using a strategy of distributing their attention over the entire area of overlapping RDKs. Such a dependence of feature selection on the regional allocation of spatial attention is consistent with some prior studies (Anllo-Vento & Hillyard, 1996; Hillyard & Münte, 1984; Karayanidis & Michie, 1996) but is more difficult to reconcile with other studies (Saenz et al., 2002; Zhang & Luck, 2009) or with the feature-similarity gain model (Maunsell & Treue, 2006). Further study is needed to resolve this question.

Acknowledgments

We would like to thank Renate Zahn, Sophie Trauer, Norman Forschack, and Christopher Gundlach for assistance with data recordings. We would also like to thank two anonymous reviewers for their comments on the manuscript. Stimulation was realized using Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience. Research was supported by the Deutsche Forschungsgemeinschaft, the U.S. National Eye Institute, and the U.S. National Institute of Mental Health.

Commercial relationships: none.
Corresponding author: Søren K. Andersen.
Email: andersen@rz.uni-leipzig.de.

References


Hillyard, S. A., & Münte, T. F. (1984). Selective attention to color and location: An analysis with event-related...


