COGNITIVE NEUROSCIENCE

A cortical zoom-in operation underlies covert shifts of visual spatial attention

Mandy V. Bartsch[1,2], Christian Merkel[1,3], Hendrik Strumpf[3], Mircea A. Schoenfeld[1,3,4], John K. Tsotsos[5,6,7,8], Jens-Max Hopf[1,3]*

Shifting the focus of attention without moving the eyes poses challenges for signal coding in visual cortex in terms of spatial resolution, signal routing, and cross-talk. Little is known how these problems are solved during focus shifts. Here, we analyze the spatiotemporal dynamic of neuromagnetic activity in human visual cortex as a function of the size and number of focus shifts in visual search. We find that large shifts elicit activity modulations progressing from highest (IT) through mid-level (V4) to lowest hierarchical levels (V1). Smaller shifts cause those modulations to start at lower levels in the hierarchy. Successive shifts involve repeated backward progressions through the hierarchy. We conclude that covert focus shifts arise from a cortical coarse-to-fine process progressing from retinotopic areas with larger toward areas with smaller receptive fields. This process localizes the target and increases the spatial resolution of selection, which resolves the above issues of cortical coding.

INTRODUCTION

Primates can intentionally direct attention to spatial locations away from where their eyes gaze. This is referred to as covert attentional focusing, an important cognitive trait that has remained a topic of intense research for now more than a century (*1*–*7*). Cell recordings in the monkey, brain imaging, and lesion studies in humans led to a detailed characterization of cortical and subcortical structures and mechanisms controlling the allocation of spatial attentional (*8*–*11*). Considerable insight was also gained from exploring visual cortical mechanisms underlying the perceptual benefits and costs associated with spatial attentional focusing (*7*, *12*–*19*).

Despite this wealth of knowledge about local cortical mechanisms and structures involved in visual spatial attention, there is still little understanding of how the hierarchy of visual cortical areas as a whole accomplishes to bring about dynamic shifts of the focus of attention in a covert way. It is important to acknowledge that covert shifts pose serious challenges for the processing across the visual cortical hierarchy (*20*, *21*) as they require a consistent recoding of target information outside foveal representations. The spatial resolution at those representations is typically low and decreases with eccentricity (*22*). Furthermore, the visual cortex is hierarchically organized (*23*, *24*), with the precision of location coding decreasing toward higher levels of representation, a loss of resolution referred to as the sampling problem (*25*). In addition, because of receptive field (RF) convergence, items falling into separate RFs at lower levels are represented by the same RF at higher levels, which leads to issues of signal cross-talk. Last, covert attention poses the problem of routing the input from different input locations toward higher-level representations (*26*). How covert focus shifts are actually accomplished in the hierarchy of visual cortex areas while simultaneously overcoming those challenges of cortical coding is unknown.

Here, we address this question using magnetoencephalography (MEG)–based source localization of cortical activity arising during covert attention shifts. We use various versions of a cued visual search task that allow us to control the degree to which attention shifts are required for target identification. As a neural measure of attention shifts, we analyze the cortical current sources of the N2pc (N2-posterior-contralateral) component, which has been shown to be a reliable index of attentional focusing in visual search (*27*). The N2pc is a well-characterized component reflecting processes of target selection, distractor attenuation (*28*, *29*), and the increase of spatial resolution within the focus of attention (*30*, *31*). The neural sources of the N2pc have been characterized in detail using MEG and electroencephalography (EEG) recordings (*15*, *30*, *32*, *33*). The N2pc allows tracking spatial shifts of attention with the highest temporal resolution (*34*, *35*).

RESULTS

Spatial focus shifts emerge from neural activity flowing top-down in the visual cortex hierarchy

To investigate spatial shifts of attention, we use versions of a cued visual search task that are sketched in Fig. 1. In experiments 1 and 2, subjects are to search for and discriminate a color-defined target (red C in Fig. 1) among four color items appearing at permanently presented placeholder positions (black squares) in the bottom left and right visual quadrant. In experiment 1, we aim to derive brain activity reflecting covert shifts of attention in search arrays as illustrated in Fig. 1A. At the beginning of a trial block, subjects are either presented with a noninformative cue (bidirectional arrow), after which the target appears randomly at any of the four placeholder positions (search condition). This requires participants to shift attention covertly (while fixating the center cross) from the center cross to the left or right visual field on each trial by 6.2° of visual angle. Alternatively, they are cued to one visual field quadrant where the target will appear (100% valid), allowing them to covertly prefocus attention (gray area) to the two placeholder positions

[1]Leibniz-Institute for Neurobiology, Magdeburg, Germany. [2]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands. [3]Otto-von-Guericke University, Magdeburg, Germany. [4]Kliniken Schmieder, Heidelberg, Germany. [5]Department of Electrical Engineering and Computer Science, York University, Toronto, Canada. [6]Centre for Innovation in Computing at Lassonde, York University, Toronto, Canada. [7]Centre for Vision Research, York University, Toronto, Canada. [8]Department of Computer Science, University of Toronto, Canada.
*Corresponding author. Email: jens-max.hopf@med.ovgu.de

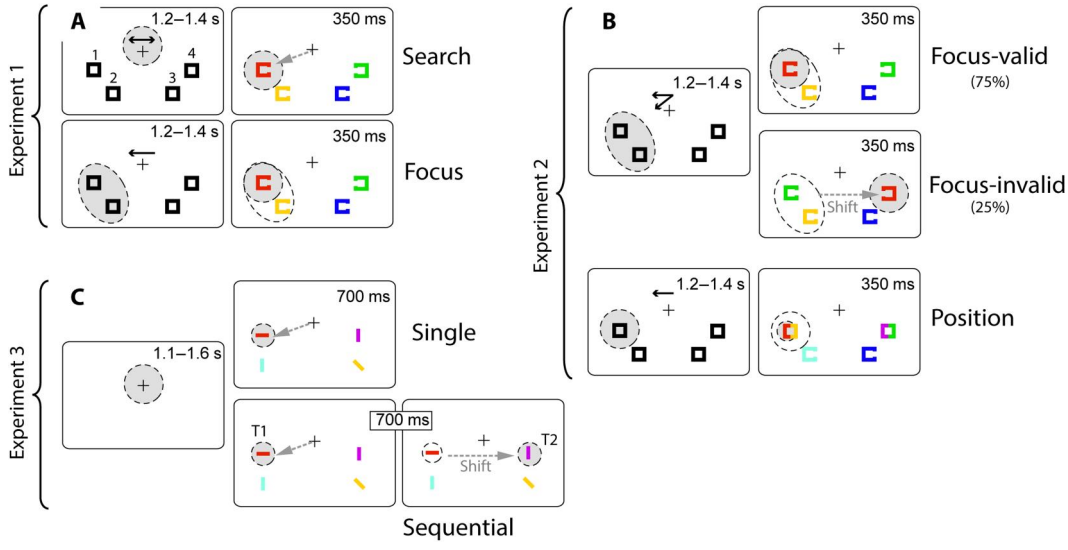

Fig. 1. Experimental design. In three experiments, participants performed different versions of a search task for a colored item, reporting its orientation ("C," site of gap; bars, vertical, diagonal, or horizontal) via button press. Gray circles illustrate the focus of spatial attention, and gray dashed arrows illustrate its shift. (**A**) Experiment 1. Permanent placeholders (black rectangles) demarcated positions of upcoming stimuli (1 to 4). Endogenous cues (black arrows) informed participants every 10 trials about the current experimental focusing condition. The cue could be uninformative (search condition) or point to the visual hemifield of the upcoming target (here, red C), allowing for prefocusing (focus condition). (**B**) Experiment 2. Focus condition as in experiment 1 (focus-valid), except that, on 25% of the trials, the target appeared in the uncued hemifield (focus-invalid). Position task (position): Target appeared at the indicated item position (one of four) together with a differently colored distractor item. At the mirror position, an analogous two-color distractor was shown. (**C**) Experiment 3. Similar to search condition of experiment 1 but without placeholders or uninformative cues. On half of the blocks, participants were instructed to report the orientation of the color-defined target (single). On the remaining trial blocks, they had to use the orientation direction of the color target (T1) to shift their attention to a second target (T2) and report its orientation (sequential). Note that the exact size and distance of the search items shown in the figure differ from how they were presented during the experiment.

(separated by 2.6°) in this quadrant (focus condition). This eliminates large shifts of the focus of attention from fixation to one or the other hemifield but leaves small focus changes (items in one quadrant), either in the form of small shifts toward one or the other position within the cued quadrant or as a zoom-in from the cued placeholders to the one containing the target. In addition, the search target is defined by color. The search, therefore, involves feature-based attention to color, which elicits cortical activity modulations similar to the N2pc (*36*). The contribution of feature-based attention, however, is controlled for because it is equally required by the search and focus conditions. Last, the experimental conditions do also not differ regarding the difficulty of the gap discriminating task. Hence, a comparison of the search versus focus condition will dissociate the part of the N2pc response that reflects large shifts of attention (from fixation to the left versus the right quadrant) from the response reflecting small focus changes within the attended quadrant.

Behavioral performance

Subjects responded slightly faster and more accurately in the focus condition [mean response time (RT) = 486 ms, correct responses = 92.0%] than in the search condition (mean RT = 496 ms, correct responses = 90.5%). This is confirmed by a repeated measures analysis of variance (rANOVA) with the two-level factor condition (search and focus), which yielded significant differences of RT [$F(1,22) = 17.9$, $P < 0.0001$] and accuracy [$F(1,22) = 12.8$, $P < 0.003$].

Neuromagnetic response

Figure 2A displays the N2pc waveforms for the search and focus conditions (black solid and dashed lines; see Materials and Methods for the derivation the waveforms and definition of sensor sites). For both conditions, the N2pc is visible as negative-polarity deflection between ~180 and 310 ms after stimulus onset (at 0 ms), which shows two maxima: a first one around 200 to 230 ms and a second around 270 to 310 ms. Figure 2B shows the topographical maps of the N2pc field distribution at 220 and 270 ms, which reveals a typical pattern with efflux-influx transition zones [red-to-blue (positive-to-negative) field lines] over occipito-temporal regions of the left hemisphere (LH) and right hemisphere (RH) (*15*, *31*, *32*). Notably, the early field transition zone of the search condition [encircled by black circle in the 220-ms map (boxed blue)] is more anterior lateral than in the focus condition (boxed red), which suggests, as confirmed by the source analysis below, that the N2pc of the search condition is initially generated at a higher level in the visual cortex hierarchy than that of the focus condition. The map at 270 ms shows that the strongest field transition of the search condition appears over more posterior regions (black ellipse, boxed turquoise), suggesting that the N2pc transitioned to lower-level areas in the visual hierarchy. The more posterior field transition is also present in the focus condition (boxed orange). To statistically validate the presence of the N2pc in both experimental conditions, a topographic analysis of variance (tANOVA) including all sensor sites was computed between 0 and 500 ms after stimulus onset (see Materials and Methods for details). The level of significance (corrected *P* values) is plotted below the waveforms in Fig. 2A, which indicates that the N2pc is significant between ~200 and 320 in both conditions. For the focus condition, there is an additional positive deflection around 160 ms (P1). This deflection represents a known modulation of the P1 component

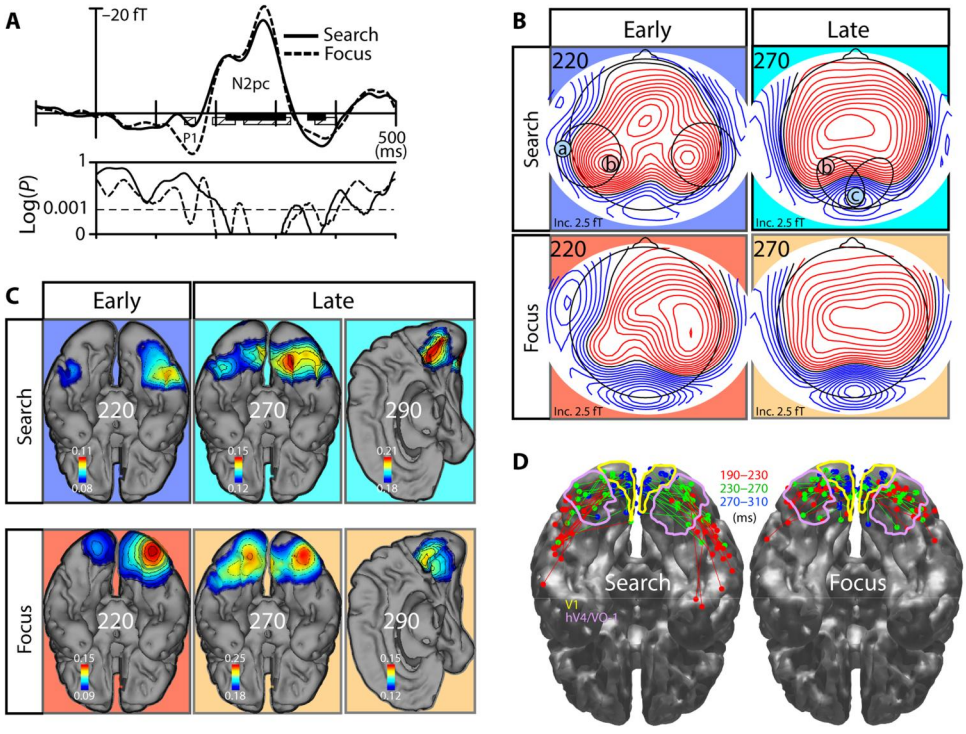

Fig. 2. Results of experiment 1. (**A**) N2pc waveforms of the search (thick solid line) and focus condition (thick dashed line). Below the waveforms, the time course of *P* values determined by tANOVAs (see Materials and Methods) testing the N2pc of the search (thin solid line) and focus (thin dashed line) conditions. The light gray horizontal dashed line marks the threshold of statistical significance. Significant time ranges in the N2pc difference are highlighted by black (search) and dashed (focus) horizontal bars below the waveforms. (**B**) Magnetic field distributions of the N2pc at early (220 ms) and late (270 ms) time points after stimulus onset. The black circles (search condition) mark the maximum efflux-influx field transition zone (red-to-blue field lines) that moves from a more anterior lateral region (difference between sensors b and a) to a posterior medial region (difference between sensors b and c). (**C**) CSD maps (views onto the basal brain) showing the N2pc sources of the search and focus conditions in early (220 ms) and late (270) time ranges corresponding to the field maps in (B) as well as at 290 ms (views at the mesial brain). The CSDs are arbitrarily thresholded (see Materials and Methods) to best illustrate the cortical localization of the source maxima. (**D**) Locations of source maxima in individual subjects in three consecutive time ranges (190 to 230 ms, red dots; 230 to 270 ms, green dots; 270 to 310 ms, blue dots) spanning the N2pc response. The red (green) lines connect corresponding 190- to 230-ms (230- to 270-ms) with the 230- to 270-ms (270- to 310-ms) maxima of individual subjects.

(*37*), which confirms that subjects prefocused attention to the side of the upcoming target in this condition. The P1 modulation is not the focus here, and a topographical analysis of this component is added in the Supplementary Materials (fig. S1).

The source activity [current source density (CSD) maps] generating the N2pc is illustrated in Fig. 2C. CSD maps (view onto the basal cortex surface) at selected time points at 220 ms (boxed blue and red) and 270 ms (boxed turquoise and orange), as well as a view onto the mesial cortex at 290 ms, are shown. A comparison of the early and late CSDs shows a transition of source maxima from higher to lower and even to lowest hierarchical levels in ventral extrastriate and striate cortex. To better evaluate the transition across hierarchical levels, Fig. 3A overlays the outlines of early and late source maxima (80% of maximum CSD) of the search (blue, early; turquoise, late) and focus (red, early; orange, late) conditions onto retinotopically defined visual areas VO-1, VO-2, hV4, and V1 (yellow and blue regions in the center of Fig. 3). Consistent with previous reports (*15*, *30*, *32*), the current sources of the N2pc appear in ventral extrastriate areas where they start at higher levels and move to lower levels of representation. In addition, as already suggested by the field distribution maps in Fig. 2B, the search condition shows initial source maxima (blue) in anterior IT regions beyond higher-level retinotopic areas VO-1/VO-2. This activity is followed by maxima around 260 to 270 ms in posterior mid-level areas hV4/VO-1 (turquoise), followed, in turn, by activity at even earlier cortical levels. This is in contrast to the focus condition, which shows first maxima (red) in mid/lower-level areas (hV4/VO-1) from where activity transitions to earlier levels (orange) including the primary visual cortex. A statistical validation of the CSD transition from higher to lower levels in the cortical hierarchy is illustrated in Fig. 2D. The cortical location of individual source maxima was determined in each subject and experimental condition (search and focus) in early (190 to 230 ms, red dots), middle (230 to 270 ms, green dots), and late (270 to 310 ms, blue dots) time ranges of the N2pc response. Each source maximum was labeled according to whether it appeared in one of three regions of interest (ROIs): IT, hV4/VO-1 (purple outline), and V1 (yellow outline). The labels were then subjected to a Friedman test with the three time ranges as factor levels, which yielded significant effects for the search and focus conditions in both hemispheres [search (LH): chi-square (2,34) = 28.11, $P < 0.0001$; search (RH): chi-square (2,34) = 31.11, $P < 0.0001$; focus (LH): chi-square (2,34) = 18.6, $P < 0.0001$; search (RH): chi-square (2,34) = 25.3, $P < 0.0001$].

The prominent difference between the search and focus conditions is that, for the former, the N2pc starts in higher-level visual

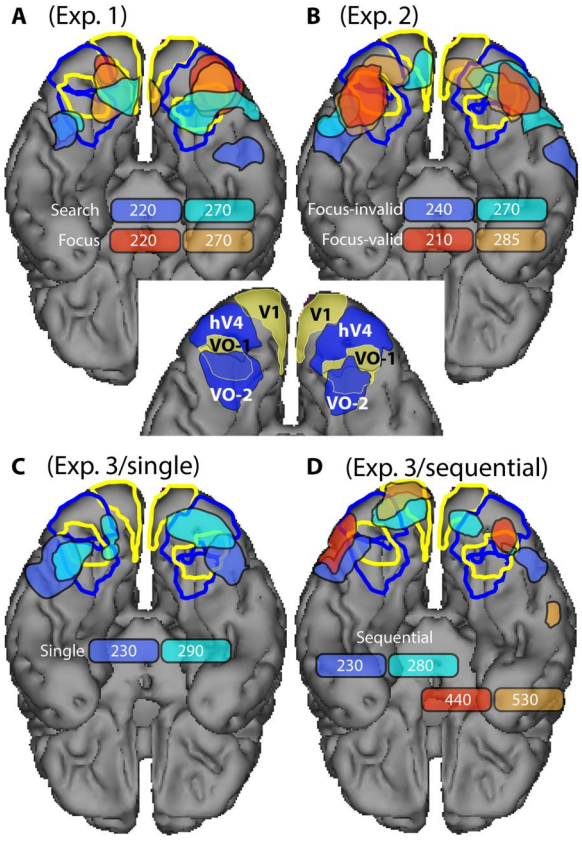

Fig. 3. Summary of CSDs across experiments. CSD distributions of the early (blue and red) and late N2pc sources (turquoise and orange) overlaid onto probabilistic maps of retinotopic visual areas (V1, hV4, VO-1, and VO-2). The latter are shown as yellow and blue areas rendered onto a cortex segmentation of the MNI152 (center). (**A** and **B**) Results of experiments 1 and 2. (**C** and **D**) Results of the single and sequential conditions of experiment 3.

cortex to subsequently appear in mid-level and lower-level areas. In the focus condition, activity starts at mid-level representations from where it progresses to lower levels. This finding aligns with a previous demonstration that an increase of the spatial resolution of target discrimination is associated with the N2pc sources appearing at progressively lower levels in the visual hierarchy (cortical coarse-to-fine process) (*30*). A narrower focus to start with in the cued visual quadrant would be coded best in a retinotopic area with higher spatial resolution. In addition, the small focus changes toward the one or the other target position would require to resolve space with higher resolution provided by the smaller RFs of mid-level and lower-level visual areas in the cortical hierarchy. For large shifts, larger RFs encompassing the more distant target positions would initially be required. Hence, source activity would start in higher-level extrastriate cortex.

To further explore this interpretation, we compared the cortical response to the target presented at one versus the other position in the same cued quadrant of the focus condition (Fig. 4). This response difference should isolate source activity reflecting the small focus changes toward the ultimate target position (gray) but eliminate activity reflecting the prefocusing to the placeholder positions (dashed ellipse), because prefocusing is equivalent for both target positions. According to our hypothesis that the size of focus change scales with the hierarchical level of representation, we predict that source activity underlying small focus changes within the attended quadrant arises in a late time range of the N2pc and at early levels in visual cortex where RFs are small enough to resolve the two target positions. Figure 4 shows that this is the case. The maxima of the response difference appear around 280 ms in early visual cortex (V1 and V2) contralateral to the cued quadrant.

To summarize, the results of experiment 1 imply that the cortical operation underlying spatial focus shifts may in essence be a cortical coarse-to-fine process, accomplished by biasing the representation of the target at progressively lower-level areas with smaller RFs, where the target can be resolved with higher resolution. This resembles a zoom-in process that locates the target by narrowing an initially wide spatial focus with low resolution down to a small one with higher resolution. This possibility is further explored in the following experiments.

Refocusing after invalid cues starts the top-down process from the highest hierarchical level

In the search condition of experiment 1, we assumed that subjects attended the fixation cross and then shifted attention to the target side. However, it is possible that, while waiting for the search frame to appear, subjects kept their attentional focus wide to encompass all four position markers before the onset of the search items. If this is the case, then the shift of attention from fixation to the target position would, in effect, include a zoom-in process from a wide to a narrow focus. Such zoom-in would be expected to arise from a coarse-to-fine process in visual cortex. Hence, to firmly demonstrate that shifts of the spatial focus per se result from a cortical coarse-to-fine process, one would have to show that such process appears when subjects shift an already narrow focus from one location to another. Experiment 2 (Fig. 1B) addresses this question with a simple modification of experiment 1. Subjects perform a cued visual search (focus condition) as in experiment 1, except that, now, the location cue is only 75% valid (focus-valid). On 25% of the trials, the target appears at one of the placeholders in the uncued visual field (focus-invalid), requiring subjects to shift their already narrow preset focus to the target in the uncued quadrant. For valid trials, we expect the N2pc sources to start at mid-level areas as in the focus condition of experiment 1. For the large focus shifts on invalidly cued trials, we would predict instead that the focused state is given up, and a coarse-to-fine process restarts from a higher-level of cortical representation (anterior IT) as in the search condition of experiment 1. An alternative possibility is that the focused state, established after the pre-cue, is simply transferred to the other visual field. In such case, the source activity underlying the N2pc on invalid trials would be expected to start at mid-level areas (hV4/VO-1) as in the focus condition of experiment 1.

A further interpretation of the results of experiment 1 based on previous work (*30*) is that a narrower preset focus would correspond with N2pc activity at lower levels of visual cortical representation, with the higher-to-lower level progression of source activity representing a zoom-in operation. To further test this interpretation, we added a third experimental condition that allowed subjects to establish an even narrower preset focus as in the focus-valid/invalid conditions. Specifically, we had trial blocks in which participants were cued to the exact placeholder position of the upcoming target (position condition), allowing them to preset the size and the resolution

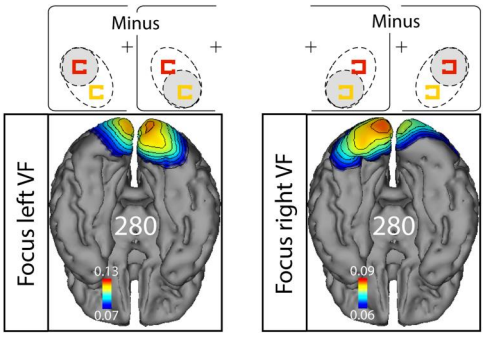

Fig. 4. Within-cued quadrant analysis (experiment 1). Left: Source estimates of the response difference elicited by a target in the top left minus the bottom right item position in the bottom left visual quadrant of the focus condition. **Right**: Response difference elicited by a target in the bottom left minus the top right item position in the bottom right visual quadrant.

of the focus accordingly. To keep the requirement of selecting the target based on color, we changed the target into a double color item (a square composed of two Cs with a target and distractor color, cf. Fig. 1B). This halved the size of the target and thus further increased the spatial resolution required for its discrimination.

Behavioral performance

As expected, subjects performed faster and more accurate in the focus-valid condition (mean RT = 550 ms, SD = 70 ms; % correct = 94.2, SD = 4.15) than in the focus-invalid condition (mean RT = 591 ms, SD = 88 ms; %correct = 90.1, SD = 10.9). Performance of the position condition was slowest and associated with intermediate accuracy (mean RT = 681 ms, SD = 92 ms; %correct = 92.1, SD = 3.7). This is consistent with the higher spatial resolution required to discriminate the target in that condition. A rANOVA with the three-level factor condition (focus-valid, focus-invalid, and position) confirmed that RT differed significantly among the experimental conditions [$F(2,38) = 24.3$, $P < 0.0001$], while accuracy did not [$F(2,38) = 2.64$, $P = 0.12$].

Neuromagnetic response

Figure 5A shows N2pc difference waveforms of the focus-valid (solid black line), focus-invalid (dashed black line), and position (gray line) conditions. As in experiment 1, the N2pc appears as negative polarity modulation between ~200 and 300 ms after the onset of the search items in the focus-valid and position condition. The modulation of the focus-invalid condition, however, is of positive polarity. This is expected, given that the focus of attention is shifted in the direction opposite to that of the focus-valid condition (see the "Derivation of the N2pc" section). Figure 5B shows CSD maps of the three experimental conditions in an early and late N2pc time range, as well as a map of the focus-invalid condition showing parietal activity before the N2pc (190 ms). The location of the source maxima of the focus-valid and focus-invalid condition in relation to the hierarchy of retinotopic areas is illustrated in Fig. 3B. Replicating the observations of experiment 1, the initial N2pc sources of the focus-valid condition (red) appear in mid-level areas (hV4/VO-1), from where they extend to earlier levels (orange). In contrast, the focus-invalid condition (blue) shows N2pc activity starting in anterior IT cortex, as in the search condition of experiment 1, which later changes to mid-level and ultimately to early levels of representation (turquoise). This clearly confirms the prediction that, on invalid trials, the coarse-to-fine process restarts at a higher-level representation than on focus-valid trials.

Notably, for the focus-invalid but not for the focus-valid condition, we see an early source maximum in left intraparietal sulcus that arises prior (at ~190 ms) to source activity in IT (Fig. 5B, boxed white). Such parietal source was previously shown to contribute to the N2pc in humans and in the monkey (*15*, *38*). It likely represents a control signal from intraparietal cortex mediating the reorienting of attention to the target item (*11*, *39*, *40*).

Last, when prefocusing the exact target position (CSD maps boxed gray in Fig. 5B), the initial N2pc sources arise at mid-lower levels as in the focus-valid condition, but they also immediately appear at the earliest level (V1) in visual cortex. This supports the interpretation that an increasingly narrow focus corresponds with N2pc activity at progressively lower levels of the visual cortical hierarchy (*30*). The position condition does not involve covert shifts of the focus of attention. It just requires target discrimination at a spatial resolution smaller than in the focus conditions, consistent with an immediate modulation of the V1, the retinotopic area with the smallest RFs and therefore the highest spatial resolution.

Sequential focus shifts involve a sequence of top-down progressions through the visual cortex hierarchy

The results of experiments 1 and 2 suggest that covert shifts of the spatial focus of attention are accomplished by a cortical zoom-in process running from higher to lower levels of representation in visual cortex. If this is the case, then one would predict that, during successive shifts, this top-down process would repeatedly appear. That is, each new shift would start over with a coarse localization at higher cortical levels to subsequently refine the localization at lower levels with increasing spatial resolution. Alternatively, the high-resolution focus established with the initial shift may be directly transferred to the retinotopic representation of the next target. This would predict that the hierarchical level of modulation does stay at the (lower) level that resolved the previous target. Experiment 3 (Fig. 1C) addresses this prediction with a sequential search task (sequential condition), where the orientation of a color-defined first target informs the subject where to focus next to discriminate a second orientation target (see Materials and Methods for details). As a control, subjects are presented with the same search arrays but are required to only search for the first target (single condition). Note that previous work has shown that sequential focus shifts are reliably indexed by a sequence of N2pc modulations (*34*, *35*).

Behavioral performance

For the single condition, RT was, on average, 632 ms, and accuracy was 93.8%. As expected, on sequential trials, the second target was reported with an average delay of 355-ms RT (987 ms) relative to the single condition and an accuracy of 89.9%. A rANOVA with the factor experimental condition (single and sequential) confirms that the RT delay [$F(1,19) = 576.8$, $P < 0.0001$] and the difference in accuracy [$F(1,19) = 7.03$, $P = 0.016$] between conditions are significant.

Neuromagnetic response

Figure 6A displays magnetic waveforms of the single and sequential conditions, in which participants were to shift from the first target in one hemifield to the second target in the opposite hemifield. In both conditions, a clear N2pc arises between ~200 to 300 ms after search frame onset (N2pc1). The sequential but not the single

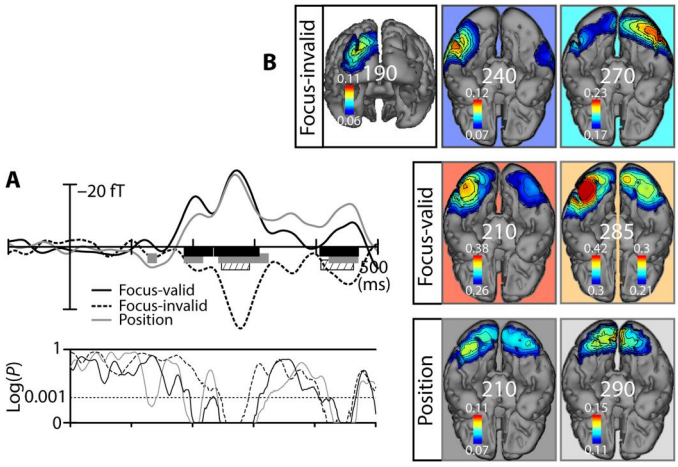

Fig. 5. Results of experiment 2. (**A**) N2pc waveforms of the focus-valid (thick solid line), focus-invalid (thick dashed line), and position (thick gray line) conditions together with the corresponding time courses of *P* values testing the N2pc for each condition. The horizontal dotted line marks the threshold of statistical significance. Significant time ranges in the N2pc difference are highlighted by black (focus-valid), dashed (focus-invalid), and gray (position) horizontal bars below the waveforms. (**B**) CSD maps showing the N2pc sources of the focus-invalid (top row), focus-valid (middle row), and position condition (bottom row) at selected time points after stimulus onset.

condition shows a second N2pc (N2pc2). To better visualize the second N2pc, the waveform averages are replotted between 300 and 700 ms relative to baseline between 300 and 400 ms after the first N2pc (gray bar). The second N2pc is highlighted by the red area under the curve that starts as positive-going modulation at ~450 ms and lasts until ~550 ms. As expected, the polarity of the second N2pc is opposite to the first one, because the second focus shift goes in the opposite horizontal direction than the first shift.

Figure 6B shows CSD maps for the single and sequential conditions. The single condition clearly reproduces the source activity underlying the N2pc of the search condition of experiment 1. Initial source activity arises in higher-level extrastriate areas (boxed blue), which is followed by activity in lower-level areas (boxed turquoise), reflecting a coarse-to-fine process in visual cortex that localizes and resolves the first target. Beyond 320 ms, no significant source activity appears in ventral extrastriate cortex. In contrast, while showing an initial coarse-to-fine process as the single condition (boxed blue and turquoise), the sequential condition is associated with a second coarse-to-fine process (boxed red and orange), starting around 450 ms after search frame onset in higher-level visual areas. From there, modulations propagate to earliest levels (V1) within ~50 ms. Figure 6C shows the time course of source activity of the single (solid line) and sequential (dashed line) conditions at the CSD maximum of the N2pc2 in early visual cortex (white circle). The second N2pc of the sequential condition is highlighted in red whose temporal maximum appears with a delay of ~210 ms relative to the first maximum. The delay can be taken as a rough estimate of the time it takes to abandon the first and establish the second focus, which is notably consistent with estimates of attentional dwell time in visual search (*41*–*43*).

Last, between 300 and 400 ms, we see source activity in left intraparietal cortex (boxed white), whose localization is very similar to the parietal source activity elicited by the focus-invalid condition seen in experiment 2 (cf. Fig. 5B). The parietal activity appears during the transition from the first to the second N2pc. This is consistent with the need to command refocusing after the first shift, which is known to rely on intraparietal cortex activity (*11*, *39*, *40*).

DISCUSSION

This study set out to investigate how covert shifts of the focus of attention are accomplished in the visual cortex hierarchy. To this end, we analyze the spatiotemporal change of current sources generating the N2pc to assess the dynamic of cortical activity underlying attentional focusing. We find that large covert shifts from fixation to the left or right visual field, as well as between target positions in opposite visual hemifields, are associated with N2pc source activity starting in high-level ventral extrastriate cortex (IT). From there, activity propagates backward through mid-level (hV4/VO-1) to lower-level areas, indicating that focus shifts involve a coarse-to-fine process progressing downward the visual cortical hierarchy. Notably, when subjects are cued to prefocus a narrower region containing two nearby target-positions, the N2pc sources start at mid-level areas. When the exact position of the target is cued, source activity appears at even earlier levels of visual representation. After invalid cues, which require shifting the focus to the opposite visual field (VF), the N2pc sources start at the level of IT. Last, when searching for two targets in sequence, with the first instructing about where to shift next, we find that each shift involves a coarse-to-fine process starting at the level of IT. The delay between both processes is found to be consistent with estimates of the attentional dwell time obtained in other experiments (*41*, *42*). This pattern of results leads to at least three conclusions.

1) In all experiments, activity modulations underlying attentional focusing progress from higher to lower levels through the visual cortex hierarchy, indicating that attentional focusing is essentially mediated by feedback processing. This conclusion dovetails with many previous studies (*19*, *31*, *32*, *44*–*47*), showing that attention effects in early visual cortex arise with a delay relative to higher-level cortex areas.

2) The narrower the region of prefocusing, the lower the hierarchical level at which the initial N2pc source activity starts to arise. This observation extends previous work showing that the locus of N2pc source activity reflects the spatial scale of target discrimination, with higher spatial resolution being associated with source maxima at progressively lower levels of representation (*30*).

3) Shifting the focus of attention involves a progression of activity from higher to lower levels in the visual hierarchy, even when a highly focused state is shifted to another location. This suggests that the attentional focus does not shift by simply transferring the neural state of being focused from one to another retinotopic representation at the same level in the cortical hierarchy. Instead, focus shifts emerge from restarting the top-down process from a level in the cortical hierarchy at which RFs are large enough to encompass the new target location. This process increases the resolution of discrimination and thereby localizes the target. In other words, the increase of spatial resolution and covert focus shifts are accomplished by the same cortical coarse-to-fine process that operates like a zoom-in operation across the visual cortex hierarchy.

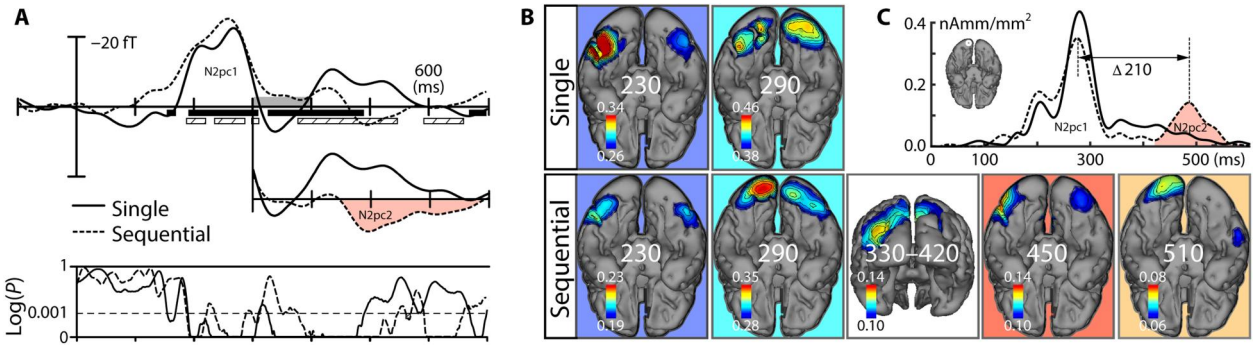

Fig. 6. Results of experiment 3. (**A**) N2pc waveforms and statistical validation of the N2pc elicited in the single (thick solid line) and sequential condition (thick dashed line). The diagram added below the waveforms replots the response of the response between 300 and 700 ms with reference to a baseline activity between 300 and 400 ms (gray horizontal bar). The second N2pc of the sequential condition (N2pc2) is highlighted as red area under the curve. (**B**) CSD maps showing the N2pc sources of the single (top row) and sequential condition (bottom row) at selected time points after stimulus onset. (**C**) Time course of the CSD estimates in an ROI in early visual cortex (white circle) of the single (solid line) and sequential condition (dashed line). Source activity reflecting the second N2pc of the sequential condition is highlighted in red.

A cortical zoom-in account of focus shifts and spatial resolution increase

Substantial behavioral work has shown that covert attention can increase spatial resolution (*48*). As an explanation, it was hypothesized that attention increases the sensitivity of neurons in low-level areas with small RFs, which would shift the net cortical sensitivity toward higher spatial frequencies and thereby increase resolution (resolution hypothesis) (*49*, *50*). Furthermore, it was proposed that the increase of spatial resolution and shifts of the focus of attention are accounted for by mechanisms of RF shift and RF shrinkage, respectively (linking hypothesis) (*48*, *51*). Consistent with the resolution and linking hypothesis, we observe that, for a narrower prefocus, the initial N2pc sources appear at a lower hierarchical level where neurons have smaller RFs. The present data suggest that both consequences of spatial focusing arise from the same coarse-to-fine process in visual cortex.

However, how exactly would focus shifts and the increase of resolution arise from modulations running from higher to lower levels in the visual cortical hierarchy? Compatible with the present observations, a number of computational models have incorporated feedback processing across the visual hierarchy as an essential attention mechanism (*52*–*54*). The present data particularly support predictions of the selective tuning (ST) model of visual attention (*25*, *52*, *55*).

The ST model implements spatial focusing as a hierarchical zoom-in process, which shifts the focus toward the target and thereby increases the resolution within the focus. Previous work has confirmed predictions of the ST model by showing that, as a direct consequence of recurrent processing in visual cortex (*46*), the focus of attention is surrounded by an inhibitory zone when target selection requires scrutiny and spatial precision (*31*, *56*). The present observations add further support to the model. In the ST model, the visual cortex hierarchy is modeled as a pyramidal structure with divergent and convergent connectivity between units at different hierarchical levels. At the top level, a unit receives many projections from input-level units, which causes large RFs and very low spatial resolution (resolution problem). To increase spatial resolution at the top, attentional focusing is implemented by a winner-take-all (WTA) process that defines a winning unit at each level of representation in reverse hierarchical direction from highest to lowest levels, thereby producing an inhibitory beam around a pass zone connecting winning units at each level. With this top-down pass through the hierarchy, the winning unit at the top progressively increases its resolution until the pass zone reaches the RF size of the input units. Now, consider location pre-cuing, where the inhibitory beam can be preset to that the pass zone selects a smaller target region before stimulus onset. Here, the WTA can start at a lower-level representation (*52*), where the RF of the winning unit represents just this region. This is what we see for the focus, the focus-valid, and the position conditions in experiments 1 and 2. Last, consider invalid cueing, where the target appears outside the preestablished beam. Because, the new target cannot be reached by a WTA within the beam (signal routing problem), the WTA process must start at a higher-level in the hierarchy, where units have projection zones large enough to contain the now attended and the new target position. This is what we see in the invalid condition of experiment 2 and the sequential condition of experiment 3. Hence, the present data strongly support the ST model's notion that the operation of shifting the beam and increasing the spatial resolution of selection emerge from the same top-down WTA process. (A more detailed discussion of the present observations in terms of the ST model is provided as Supplementary Text and fig. S2.)

Resolving challenges of target coding during covert focus shifts

As outlined in Introduction, covert shifts face challenges of cortical recoding in terms of low spatial resolution, cross-talk, and signal routing due to the convergent forward hierarchy of the visual cortex (*20*, *25*, *26*). The here identified top-down coarse-to-fine process underlying both, focus shifts and the resolution of attention, provides a direct solution to the challenges of spatial resolution and signal routing. Moreover, it endorses an implementation as suggested by the ST model, which resolves issues of cross-talk and ambiguous coding (*25*), and provides a front-end of selection for fixation in overt focusing (*55*). As the top-down process progresses to lower levels in the hierarchy, forward projections from nontarget units are more and more excluded from contributing to the RF of top-level units coding the target, which disambiguates the representation of the target, i.e., eliminates cross-talk.

Last, the present data account for some observations that contradicted the view that the focus of attention moves in an analog manner across space (*3*, *57*). The time to shift to a cued location was found to be independent of the distance to the current focus, suggesting discrete jumps of the focus (*58*–*60*). Also, symbolic location cues created initial benefits at all locations in a hemifield followed by a loss of benefits at uncued locations (*5*), suggesting that an initial broad beam is narrowed down to the target location instead of moving in a focused state across the visual field. Those observations are easily accounted for by the present findings, which suggest that focus shifts are brought about by a cortical zoom-in process in the visual hierarchy where the time required to shift the focus may rather depend on the hierarchical distance in visual cortex than the spatial distance between items in the visual scene. That is, as long as the coarse-to-fine process starts and ends at the same hierarchical level, the time to establish the attentional beam would be the same for a more or less distant target location.

MATERIALS AND METHODS

Participants

Twenty-three participants took part in experiment 1 (8 females; mean age, 26.8 years), 20 participants took part in experiment 2 (10 females; mean age, 27.2), and 20 participants took part in experiment 3 (9 females; mean age, 26.8). All participants had normal or corrected-to-normal visual acuity. They gave written informed consent before the measurement and were paid for participation (8€/hour, 2 to 3 hours per session including preparation). The experiments were conducted in accordance to the research regulations of the Declaration of Helsinki, and all procedures and methods were approved by the ethics board of the Otto-von-Guericke University of Magdeburg. The number of trials per experimental condition and the number of participants were based on previous work (*32*) and on recommendations for medium-sized EEG/MEG components (*61*).

Experimental design

Experiment 1

Stimuli. As shown in Fig. 1A, search frames were bilateral consisting of four colored rectangular Cs (two in the left and two in the right visual field) presented at permanently highlighted placeholder positions (black squares). The gap of the colored Cs was randomly oriented to the left or right. Each search display contained one red C and one green C in opposite VFs. The color of the remaining two Cs was randomly chosen from a set of six other colors (blue, yellow, cyan, magenta, purple, and orange) with the restriction that all of these colors were equally often used and all items of a search display were drawn in a different color. Before the experiment, all colors were psychophysically matched in a subset of five observers using heterochromatic flicker photometry. The search items and placeholders appeared on a gray background with a luminance of 24.6 Cd/m^2 (experiments 1 and 2) or 15.3 Cd/m^2 (experiment 3). The Cs (width, 1° visual angle; height, 1.3°) were presented in the bottom left and right visual field at polar angle positions of 31°, 56°, 124°, and 149° on an imagined iso-eccentric circle subtending a distance of 6.2° from fixation. The distance between items in one quadrant was 2.6° of visual angle. The placeholders were black squares (3.9 Cd/m^2) that exactly occupied the positions of the upcoming Cs, such that, at search frame onset, the placeholders were entirely replaced by the search items. White arrow cues (76.5 Cd/m^2; length, 2.2°; height, 1°) were shown every 10 trials centrally and 0.8° above the fixation cross.

Procedure. Participants fixated the central cross (always present) while searching for a color-defined target (either the red or green rectangular C). The target color alternated across experimental blocks and was assigned at the beginning of each block. The task was to report the gap orientation of the target C with a button press of their right hand (left, index finger; right, middle finger). The experimental conditions were run in short microblocks of 10 trials: Every 10 trials, an arrow cue appeared that could either be uninformative with respect to the target location (bilateral arrow, search condition) or indicate in which hemifield the target would appear (unilateral arrow, focus condition, 100% valid). Microblocks were used instead of mixing uninformative and informative cues on a trial-to-trial basis to minimize demands on decoding the cue and to facilitate consistent prefocusing after informative cues. All conditions were pseudorandomized such that two repetitions of the search condition (in total 20 trials) were always alternated with two focus conditions (10 trials focus left and 10 trials focus right). Having participants perform the same number of trials for the search and focus conditions was important, to gain the same number of trials with the target appearing in the left and right visual field for both the focus and search condition. The cue was shown every 10 trials for 2.5 s together with the placeholders, followed by the presentation of the search array that replaced the placeholders for 350 ms, after which placeholders reappeared. The next trial started after a variable interstimulus interval of 1.2 to 1.4 s. Each block was about 5 min long. Participants performed a total of 12 blocks resulting in a total of 1915 trials (about 478 to 480 trials for each of the four conditions: focus, left/right; search, left/right).

Placeholders. To guarantee consistent prefocusing in the focus condition of experiment 1 (and experiment 2), we permanently marked the location of the upcoming items with location placeholders that were replaced by the onset of the search items. It has been shown in macaques that the attention-related neural responses to spatial (but not feature) cues disappear in area V4 when emptying the displays after cue offset (*62*). This suggested that spatial attentional biases to symbolic cues are not consistently maintained and require a fresh rebuilt upon stimulus onset to become effective. To avoid "spurious" shifts of attention due to such breakdown of the cuing state, we use placeholders permanently highlighting the item positions. Note that a previous N2pc study of cued visual search showed that placeholders do effectively anchor the spatial focus before search frame onset (*63*).

Experiment 2

The stimulus design of experiment 2 is depicted in Fig. 1B. Trial structure and timing as well as block structure were identical to the focus condition of experiment 1 with the following exceptions: (i) For the focus condition (six blocks), the cue was only 75% valid. On valid trials (focus-valid), subjects would perform as in the focus condition of experiment 1. After invalid (cues 25%) trials (focus-invalid), the target item appeared in the opposite, uncued hemifield. On those trials, participants had to shift their spatial focus across the visual field. (ii) A new focusing condition was introduced in which an arrow cue informed about the exact position of the upcoming target (position condition). Here, the target was one of two opposing Cs with different color that both occupied the cued placeholder position. Despite knowing the exact target position, subjects still

had to select the target based on color. To balance color distribution across the screen (and be able to derive the N2pc from identical stimulus displays), an analogous compound item containing the other potential target color (red or green) was shown at the mirror position in the opposite hemifield. The position condition was run after the focus conditions in the last six experimental blocks. Here, cues were again 100% valid. (iii) The cueing procedure had to be changed relative to experiment 1. For the position condition to indicate the top or bottom item position in the target quadrant, a single arrow cue was pointing horizontally (top placeholder) or 45° downward (bottom placeholder) to the target quadrant, respectively. For the focus conditions, two arrows, one horizontal and one pointing 45° downward to the target quadrant, were presented. Analogous to experiment 1, cues were presented every 10 trials, and cued visual fields/positions were randomly alternated. Target color (red or green) was alternated between experimental blocks. Subjects performed a total of 12 experimental blocks, amounting to a total of 1920 trials (960 trials for the position task including 240 for each position: 360, focus left; 360, focus right; and 240, invalid focus).

Experiment 3

Stimuli. Stimuli and trial structure are shown in Fig. 1C. The general layout (bilateral search display with four items) was similar to experiments 1 and 2. Stimulus locations were slightly changed such that the search items were at the corners of an imagined rectangle: All items were presented 5.25° lateral to the fixation cross and 1.85° (top items) or 5.16° (bottom items) below the fixation cross. The search items were now small bars (length, 0.57°; width, 0.23°), randomly appearing with a horizontal, vertical, or a 45° radially tilted orientation. These stimulus changes were introduced for the sequential condition to have the initial target "point" to another item position depending on its orientation (e.g., a top left horizontal bar would point to the item presented at the mirror-image position in the opposite visual field). One item was always red or green, and the color of the remaining three items was randomly drawn from the six additional colors (blue, yellow, cyan, magenta, purple, and orange). No cues or placeholders were displayed.

Procedure. For six experimental blocks, participants performed a search task similar to the search condition of experiment 1 (single condition). That is, they searched for a color-defined item (red or green, blockwise alternating) and reported its orientation via button press with the right hand (vertical, index finger; diagonal, middle finger; and horizontal, ring finger). On the remaining six blocks, participants were informed that they would now have to follow the orientation of the search item (T1) to find a second item (T2) of which the orientation had then to be reported (sequential condition). As depicted in Fig. 1C, T1 could, e.g., point to a target in the opposite hemifield (vertical or diagonal), requiring participants to shift their attention between visual hemifields. A vertical T1 would require to shift to the other item in the same VF. To give participants enough time for the sequential task and keep conditions physically identical, the search frame was always shown for 700 ms with a variable interstimulus interval (ISI) of 1.1 to 1.6 s. Participants performed a total of 12 experimental blocks (each block about 4.7 min), amounting to a total of 1496 trials [about 748 trials for single, 499 trials for intervisual field shifts (left/right), and 249 trials for within-visual field shifts (up/down, not analyzed here)].

Data recording

The MEG data were continuously recorded with an Elekta Neuromag TRIUX triple sensor system including 102 magnetometers (Elekta Neuromag Oy, Helsinki, Finland). Participants were seated below the MEG dewar in a magnetically shielded chamber. Stimulus presentation was coordinated with MATLAB (MathWorks Inc., Natick, MA, USA), and the stimuli were back-projected by a DLP-LED PROPixx projector (VPixx Technologies Inc., Saint-Bruno, QC, Canada) onto a semitransparent screen (VPixx) placed in 1-m viewing distance to the participants. Participants gave responses with the right hand via a LUMItouch response box (PhotonControl Inc., Burnaby, DC, Canada) for experiment 1 and RESPONSEPixx response boxes (VPixx) for experiments 2 and 3. The head position within the MEG dewar was determined by digitizing anatomical landmarks (nasion and bilateral preauricular points) and five spatially distributed coils (near vertex, inion, nasion, and preauricular points) with a Polhemus system (Polhemus 3Space Fastrak system). Data were sampled at 1000 Hz and band-pass–filtered online from 0.03 to 330 Hz. Signal space separation provided with the Elekta Neuromag Data Analysis Software MaxFilter was used offline for the suppression of environmental noise and spatial interferences. In addition, slight changes of the head position were tracked (coil measurements after each experimental block) and used to correct for head movements during the measurement by spatially realigning the data of each subject using MaxFilter.

EEG data were simultaneously recorded using a Neuroscan system (NeuroScan, El Paso, TX, RRID:SCR_015818) and a 32-electrode cap with mounted sintered Ag/AgCl electrodes (Easycap, Herrsching, Germany), the right mastoid was used as online reference. To record the EOG (electrooculogram), one electrode was placed below the right eye (vertical EOG), and two electrodes were placed at the outer canthi of both eyes (horizontal EOG). Contact to head surface was established using an abrasive gel (Abralyt light, Easycap), all impedances were kept below 5 kilohm. The EOG was used to track eye movements. To verify the quality of fixation, a detailed analysis of the horizontal electrooculogram (HEOG) response in experiment 1 together with a control experiment deriving a reference HEOG response when actively foveating the placeholder positions is provided in the Supplementary Materials (fig. S3). The EEG data are not reported here.

Data analysis and statistical validation

After noise cancellation and the realignment of head positions (see above), the data were epoched (500 ms before to 1500 ms after stimulus onset) and further processed using MATLAB and the Fieldtrip toolbox (RRID: SCR_004849) (*64*). Only trials with correct responses were analyzed. In particular, data were offline-filtered (100-Hz low pass, 50-Hz notch) and detrended, and an artifact rejection was performed with thresholds being individually adjusted for each participant in the following way. All epochs in which peak-to-peak amplitudes in a time window of 100 ms exceeded an individually set threshold (EOG, maximum of 120 μV; MEG, maximum of 4 pT) were rejected, such that epochs containing eye blinks and eye movements breaking fixation (overall 15 to 25% of the epochs) were eliminated. After artifact rejection, trials were averaged for every experimental condition within participants and afterward across participants. The resulting grand average data shown in the

manuscript were baselined to a period of −200 to 0 ms before search frame onset.

Derivation of the N2pc

The N2pc is derived in a standard way by subtracting the brain response elicited by the same search array where the focus of attention was contralateral minus ipsilateral to a recording electrode or sensor [cf. (*27*, *32*)]. This subtraction will eliminate basic stimulus-driven activity and all nonlateralized attention-driven responses. In the magnetic field distribution of magnetometer data, the N2pc typically appears as mirror-image efflux-influx configuration [positive (red)–to–negative (blue) field line transition] in the LH and RH (encircled in black in Fig. 2B, top row), with the locus of strongest field transition typically reflecting the underlying source maximum. Figure 2B illustrates the derivation of the N2pc waveforms (displayed in Fig. 2A) for the early and late transition zones in the LH. The transition zones are highlighted by black circles and ellipses, respectively, in the LH and RH. For example, we derived the waveform of the anterior efflux-influx effect by averaging the influx response at sensor a with the corresponding polarity inverted efflux response at sensor b (efflux). The same was done for the posterior efflux-influx effect (average of c and b). Note that which component of the flux-field is polarity reversed is arbitrary. We decided to reverse the positive field component (efflux), which derives the magnetic analog of the N2pc as a negative polarity deflection in the ERMF response. This was solely done to keep terminology consistent with the N2pc of the ERP. To simplify data presentation, the waveforms of the anterior and posterior transition zone, as well of the LH and RH, were averaged. Note that the polarity of the N2pc depends on the direction of subtraction. Subtracting an ipsilateral from a contralateral target, for example, in the invalid condition of experiment 2, will flip the N2pc to opposite polarity.

Statistical validation

A tANOVA as implemented in the multimodal neuroimaging software Curry 8 was used to validate the N2pc component statistically. Specifically, for each experiment and experimental condition, a one-way tANOVA (*65*) with the factor target visual field (left and right) was computed in a time range between 0 and 500 ms (experiments 1 and 2) or between 0 and 700 ms (experiment 3) after stimulus onset. All sensors (102 magnetometers) were taken into account. Before the statistical analyses, the data were low-pass–filtered with a cutoff at 50 Hz. The tANOVA is based on nonparametric permutation performed for each time sample in the time interval of interest. The resulting issue of multiple comparisons in the time domain was dealt with using spectral parameters of the data as suggested by Wagner and colleagues (*66*). For the present calculations, a sampling frequency of 500 Hz and a filter frequency of 50 Hz were used, which resulted in a corrected P value of 0.001 (horizontal dashed line in Figs. 2A, 5A, and 6A). The sensor sites, the N2pc waveforms are shown from in those figures (a, b, and c in Fig. 2B; and mirror-image positions in the RH), were defined on the basis of the most significant entries in the channel impact maps of the tANOVAs (*67*) in the time range of the N2pc (200 to 300 ms). The channel impact map returns the degree to which a sensor reflects (and therefore illustrates) the difference under consideration (the N2pc effect) in the tANOVA.

Source localization analysis

Source localization analysis (CSD analysis) was performed using the minimum least-squares method as implemented in the multimodal neuroimaging software Curry 8 (Compumedics USA Inc.). We used a fixed orientation (surface normal) model with a spatial smoothing of 5 mm. The regularization parameter was determined with the $\chi 2$ criterion, assuming that $\Delta 2$ should be approximately in the range of data noise. As we analyze data averaged over subjects, a segmentation of the MNI152 brain (standard brain, provided in Curry 8) served as source compartment. The MNI brain represents an anatomical average across 152 observers, which should reasonably accommodate the magnetic field data smoothing coming with across subject averages. Noise estimates were derived from a baseline 200 ms before search frame onset. CSDs were estimated for each time sample in a time range between 0 and 500 ms after search frame onset. For visualization in the figures, the CSD data were smoothed with a sliding average filter (window width, 20 ms). To illustrate source maxima, the CSD maps shown in Figs. 2, 4, 5, and 6 were set to an arbitrary threshold between 65 and 75% of the source density maximum at a given time point after stimulus onset. To facilitate the visualization of the source maxima in Fig. 3, the outlines of the shown CSD distributions were set to ~80% of the CSD maximum independently in each region displaying a maximum in the given time range.

Statistical validation of the propagation of source maxima from higher to lower levels in the cortical hierarchy in experiment 1 was performed using rank order testing (Friedman test). To this end, CSD maps were estimated for each subject and experimental condition using the standard MNI cortical surface segmentation. Source maxima for the search and focus conditions were determined in three subsequent 40-ms time windows spanning the time range of the N2pc (190 to 230, 230 to 270, and 270 to 310 ms), separately for the LH and RH. Each maximum was then assigned a label which indicated in which of the three ROIs the maximum appeared: IT, hV4/VO-1, and V1. A Friedman test with the three time ranges serving as factor levels was lastly performed for the search and focus conditions, testing the null hypothesis that ROI labels would be randomly assigned in the three time ranges.

Retinotopic maps

The retinotopic areas displayed in Fig. 3 are probabilistic maps (*68*, *69*) downloadable from https://scholar.princeton.edu/napl/resources. The maps were rendered onto a 1-mm surface segmentation of the MNI152 brain using FreeSurfer (version 5.1). The spatial extension thresholds of the shown areas (V1, hV4, VO-1, and VO-2) were set to 40%.

Behavioral data

Response time and accuracy were calculated using MATLAB (MathWorks Inc., Natick, MA, USA). Response times were calculated for correct trials only, and the effects of response time and accuracy were statistically validated using one-way rANOVAs in SPSS (SPSS Inc., Chicago, IL, USA). The significance level was set to $\alpha = 0.05$.

Supplementary Materials

This PDF file includes:

Supplementary Text

Figs. S1 to S3
References (*70*–*73*)

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. H. von Helmholtz, *Handbuch der Physiologischen Optik* (Verlag von Leopold Voss, 1896).
2. M. I. Posner, Orienting of attention. *Q. J. Exp. Psychol.* **32**, 3–25 (1980).
3. G. L. Shulman, R. W. Remington, J. P. McLean, Moving attention through visual space. *J. Exp. Psychol. Hum. Percept. Perform.* **5**, 522–526 (1979).
4. C. W. Eriksen, Y.-Y. Yeh, Allocation of attention in the visual field. *J. Exp. Psychol. Hum. Percept. Perform.* **11**, 583–597 (1985).
5. M. Shepherd, H. J. Müller, Movement versus focusing of visual attention. *Percept. Psychophys.* **46**, 146–154 (1989).
6. D. LaBerge, R. L. Carlson, J. K. Williams, B. G. Bunney, Shifting attention in visual space: Tests of moving-spotlight models versus an activity-distribution model. *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 1380–1392 (1997).
7. M. Carrasco, Visual attention: The past 25 years. *Vision Res.* **51**, 1484–1525 (2011).
8. J. B. Hopfinger, M. H. Buonocore, G. R. Mangun, The neural mechanisms of top-down attentional control. *Nat. Neurosci.* **3**, 284–291 (2000).
9. M. Corbetta, G. L. Shulman, Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
10. S. E. Petersen, M. I. Posner, The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* **35**, 73–89 (2012).
11. J. W. Bisley, M. E. Goldberg, Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.* **33**, 1–21 (2010).
12. R. Desimone, J. Duncan, Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
13. S. A. Hillyard, G. R. Mangun, Sensory gating as a physiological mechanism for visual selective attention, in *Current Trends in Event-Related Potential Research*, R. Johnson, J. W. Rohrbaugh, R. Parasuraman, Eds. (Elsevier Science Publishers B.V., 1987), pp. 61–67.
14. J. A. Brefczynski, E. A. DeYoe, A physiological correlate of the "spotlight" of visual attention. *Nat. Neurosci.* **2**, 370–374 (1999).
15. J.-M. Hopf, S. J. Luck, M. Girelli, T. Hagner, G. R. Mangun, H. Scheich, H.-J. Heinze, Neural sources of focused attention in visual search. *Cereb. Cortex* **10**, 1233–1241 (2000).
16. S. Kastner, L. G. Ungerleider, Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* **23**, 315–341 (2000).
17. J. H. R. Maunsell, Neuronal mechanisms of visual attention. *Annu. Rev. Vis. Sci.* **1**, 373–391 (2015).
18. J.-M. Hopf, H. J. Heinze, M. A. Schoenfeld, S. A. Hillyard, Spatio-temporal analysis of visual attention, in *The Cognitive Neurosciences IV*, M. S. Gazzaniga, Ed. (MIT Press, 2009), pp. 235–250.
19. T. Noesselt, S. A. Hillyard, M. G. Woldorff, A. Schoenfeld, T. Hagner, L. Jancke, C. Tempelmann, H. Hinrichs, H.-J. Heinze, Delayed striate cortical activation during spatial attention. *Neuron* **35**, 575–587 (2002).
20. C. H. Anderson, D. C. Van Essen, Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6297–6301 (1987).
21. J. K. Tsotsos, Analyzing vision at the complexity level. *Behav. Brain Sci.* **13**, 423–469 (1990).
22. G. Westheimer, The spatial grain of the perifoveal visual field. *Vision Res.* **22**, 157–162 (1982).
23. D. J. Felleman, D. C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
24. D. J. Kravitz, K. S. Saleem, C. I. Baker, L. G. Ungerleider, M. Mishkin, The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**, 26–49 (2013).
25. J. K. Tsotsos, *A Computational Perspective on Visual Attention* (MIT Press, 2011).
26. B. A. Olshausen, C. A. Anderson, D. C. Van Essen, A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J. Neurosci.* **13**, 4700–4719 (1993).
27. S. J. Luck, Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components, in *Oxford Handbook of Event-Related Potential Components*, S. J. Luck, E. Kappenman, Eds. (Oxford Univ. Press, 2011).
28. S. J. Luck, M. Girelli, M. T. McDermott, M. A. Ford, Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention. *Cogn. Psychol.* **33**, 64–87 (1997).
29. C. Hickey, V. Di Lollo, J. J. McDonald, Electrophysiological indices of target and distractor processing in visual search. *J. Cogn. Neurosci.* **21**, 760–775 (2009).
30. J.-M. Hopf, S. J. Luck, K. Boelmans, A. Schoenfeld, N. Boehler, J. W. Rieger, H.-J. Heinze, The neural site of attention matches the spatial scale of perception. *J. Neurosci.* **26**, 3532–3540 (2006).
31. C. N. Boehler, J. K. Tsotsos, M. A. Schoenfeld, H.-J. Heinze, J.-M. Hopf, Neural mechanisms of surround attenuation and distractor competition in visual search. *J. Neurosci.* **31**, 5213–5224 (2011).
32. S. E. Donohue, M. A. Schoenfeld, J.-M. Hopf, Parallel fast and slow recurrent cortical processing mediates target and distractor selection in visual search. *Commun. Biol.* **3**, 689 (2020).
33. D. Tay, D. L. McIntyre, J. J. McDonald, Searching for visual singletons without a feature to guide attention. *J. Cogn. Neurosci.* **34**, 2127–2143 (2022).
34. G. F. Woodman, S. J. Luck, Electrophysiological measurement of rapid shifts of attention during visual search. *Nature* **400**, 867–869 (1999).
35. G. F. Woodman, S. J. Luck, Serial deployment of attention during visual search. *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 121–138 (2003).
36. M. Bartsch, C. N. Boehler, C. Stoppel, C. Merkel, H. J. Heinze, M. A. Schoenfeld, J. M. Hopf, Determinants of global color-based selection in human visual cortex. *Cereb. Cortex* **25**, 2828–2841 (2015).
37. G. R. Mangun, Neural mechanisms of visual selective attention. *Psychophysiology* **32**, 4–18 (1995).
38. B. Herrera, J. A. Westerberg, M. S. Schall, A. Maier, G. F. Woodman, J. D. Schall, J. J. Riera, Resolving the mesoscopic missing link: Biophysical modeling of EEG from cortical columns in primates. *Neuroimage* **263**, 119593 (2022).
39. S. Yantis, J. Schwarzbach, J. T. Serences, R. L. Carlson, M. A. Steinmetz, J. J. Pekar, S. M. Courtney, Transient neural activity in human parietal cortex during spatial attention shifts. *Nat. Neurosci.* **5**, 995–1002 (2002).
40. J. Gottlieb, From thought to action: The parietal cortex as a bridge between perception, action, and cognition. *Neuron* **53**, 9–16 (2007).
41. C. M. Moore, H. Egeth, L. Berglan, S. J. Luck, Are attentional dwell times inconsistent with serial visual search? *Psychon. Bull. Rev.* **3**, 360–365 (1996).
42. J. Theeuwes, R. Godijn, J. Pratt, A new estimation of the duration of attentional dwell time. *Psychon. Bull. Rev.* **11**, 60–64 (2004).
43. A. Petersen, S. Kyllingsbæk, C. Bundesen, Measuring and modeling attentional dwell time. *Psychon. Bull. Rev.* **19**, 1029–1046 (2012).
44. A. Martinez, L. Anllo-Vento, M. I. Sereno, L. R. Frank, R. B. Buxton, D. J. Dubowitz, E. C. Wong, H. Hinrichs, H. J. Heinze, S. A. Hillyard, Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat. Neurosci.* **2**, 364–369 (1999).
45. A. D. Mehta, I. Ulbert, C. E. Schroeder, Intermodal selective attention in monkeys. I: Distribution and timing of effects across visual areas. *Cereb. Cortex* **10**, 343–358 (2000).
46. C. N. Boehler, J. K. Tsotsos, A. Schoenfeld, H.-J. Heinze, J.-M. Hopf, The center-surround profile of the focus of attention arises from recurrent processing in visual cortex. *Cereb. Cortex* **19**, 982–991 (2009).
47. E. A. Buffalo, P. Fries, R. Landman, H. Liang, R. Desimone, A backward progression of attentional effects in the ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 361–365 (2010).
48. K. Anton-Erxleben, M. Carrasco, Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nat. Rev. Neurosci.* **14**, 188–200 (2013).
49. Y. Yeshurun, M. Carrasco, The locus of attentional effects in texture segmentation. *Nat. Neurosci.* **3**, 622–627 (2000).
50. Y. Yeshurun, M. Carrasco, Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* **396**, 72–75 (1998).
51. M. Carrasco, A. Barbot, How attention affects spatial resolution. *Cold Spring Harb. Symp. Quant. Biol.* **79**, 149–160 (2014).
52. J. K. Tsotsos, S. M. Culhane, W. Y. K. Wai, Y. Lai, N. Davis, F. Nuflo, Modeling visual attention via selective tuning. *Artif. Intell.* **78**, 507–545 (1995).
53. G. Deco, J. Zihl, A neurodynamical model of visual attention: Feedback enhancement of spatial resolution in a hierarchical system. *J. Comput. Neurosci.* **10**, 231–253 (2001).
54. T. Miconi, R. VanRullen, A feedback model of attention explains the diverse effects of attention on neural firing rates and receptive field structure. *PLOS Comput. Biol.* **12**, e1004770 (2016).
55. J. Tsotsos, I. Kotseruba, C. Wloka, A focus on selection for fixation. *JEMR.* **9**, 10.16910/jemr.9.5.2, (2016).
56. J.-M. Hopf, C. N. Boehler, S. J. Luck, J. K. Tsotsos, H.-J. Heinze, M. A. Schoenfeld, Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1053–1058 (2006).
57. Y. Tsal, Movements of attention across the visual field. *J. Exp. Psychol. Hum. Percept. Perform.* **9**, 523–530 (1983).
58. R. Remington, L. Pierce, Moving attention: Evidence for time-invariant shifts of visual selective attention. *Percept. Psychophys.* **35**, 393–399 (1984).

59. D. Sagi, B. Julesz, Fast noninertial shifts of attention. *Spat. Vis.* **1**, 141–149 (1985).
60. J. D. Golomb, A. C. Marino, M. M. Chun, J. A. Mazer, Attention doesn't slide: Spatiotopic updating after eye movements instantiates a new, discrete attentional locus. *Atten. Percept. Psychophys.* **73**, 7–14 (2011).
61. S. J. Luck, *An Introduction to the Event-Related Potential Technique* (MIT Press, 2005).
62. B. Y. Hayden, J. L. Gallant, Time course of attention reveals different mechanisms for spatial and feature-based attention in area v4. *Neuron* **47**, 637–643 (2005).
63. G. Woodman, J. Arita, S. Luck, A cuing study of the N2pc component: An index of attentional deployment to objects rather than spatial locations. *Brain Res.* **1297**, 101–111 (2009).
64. R. Oostenveld, P. Fries, E. Maris, J. M. Schoffelen, FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).
65. T. Koenig, L. Melie-García, A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topogr.* **23**, 233–242 (2010).
66. M. Wagner, C. Ponton, R. Tech, M. Fuchs, J. Kastner, Non-parametric statistical analysis of EEG/MEG map topographies and source distributions on the epoch level. *Kognitive Neurophysiologie des Menschen.* **7**, 1–23 (2014).
67. M. Wagner, R. Tech, M. Fuchs, J. Kastner, F. Gasca, Statistical non-parametric mapping in sensor space. *Biomed. Eng. Lett.* **7**, 193–203 (2017).
68. M. J. Arcaro, S. A. McMains, B. D. Singer, S. Kastner, Retinotopic organization of human ventral visual cortex. *J. Neurosci.* **29**, 10638–10652 (2009).
69. L. Wang, R. E. Mruczek, M. J. Arcaro, S. Kastner, Probabilistic maps of visual topography in human cortex. *Cereb. Cortex* **25**, 3911–3931 (2015).
70. H. J. Heinze, S. J. Luck, G. R. Mangun, S. A. Hillyard, Visual event-related potentials index focussed attention to bilateral stimulus arrays: I. Evidence for early selection. *Electroencephalogr. Clin. Neurophysiol.* **75**, 511–527 (1990).
71. G. R. Mangun, S. A. Hillyard, Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 1057–1074 (1991).
72. H. J. Heinze, G. R. Mangun, W. Burchert, H. Hinrichs, M. Scholz, T. F. Münte, A. Gös, M. Scherg, S. Johannes, H. Hundeshagen, M. S. Gazzaniga, S. A. Hillyard, Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* **372**, 543–546 (1994).
73. G. R. Mangun, J. Hopfinger, C. Kussmaul, E. Fletcher, H. J. Heinze, Covariations in ERP and PET measures of spatial selective attention in human extrastriate visual cortex. *Hum. Brain Mapp.* **5**, 273–279 (1997).

Acknowledgments: We thank C. Teichmann for collecting the data of experiment 1. **Funding:** This work was supported by DFG grant Ho 1965/2-1 (J.-M.H.) and DFG CRC 1436 TP B5 (J.-M.H. and M.A.S.). **Author contributions:** Conceptualization: M.V.B., J.K.T., and J.-M.H. Methodology: M.V.B., H.S., C.M., and J.-M.H. Investigation: M.V.B. and H.S. Visualization: M.V.B. and J.-M.H. Supervision: M.A.S., J.K.T., and J.M.H. Writing—original draft: J.-M.H. and M.V.B. Writing—review and editing: J.-M.H., J.K.T., and M.V.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Behavioral performance data and data underlying the source statistics of experiment 1 are accessible at OSF DOI 10.17605/OSF.IO/9EJ2A.

Submitted 8 September 2022
Accepted 7 February 2023
Published 8 March 2023
10.1126/sciadv.ade7996