

Alpha-band Activity Tracks the Zoom Lens of Attention

Tobias Feldmann-Wüstefeld¹ and Edward Awh²

Abstract

■ Voluntary control over spatial attention has been likened to the operation of a zoom lens, such that processing quality declines as the size of the attended region increases, with a gradient of performance that peaks at the center of the selected area. Although concurrent changes in activity in visual regions suggest that zoom lens adjustments influence perceptual stages of processing, extant work has not distinguished between changes in the spatial selectivity of attention-driven neural activity and baseline shift of activity that can increase mean levels of activity without changes in selectivity. Here, we distinguished between these alternatives by

measuring EEG activity in humans to track preparatory changes in alpha activity that indexed the precise topography of attention across the possible target positions. We observed increased spatial selectivity in alpha activity when observers voluntarily directed attention toward a narrower region of space, a pattern that was mirrored in target discrimination accuracy. Thus, alpha activity tracks both the centroid and spatial extent of covert spatial attention before the onset of the target display, lending support to the hypothesis that narrowing the zoom lens of attention shapes the initial encoding of sensory information. ■

INTRODUCTION

Focusing visual attention improves processing at locations within the attended region (Carrasco, 2011; Posner, Snyder, & Davidson, 1980; Shaw & Shaw, 1977). A long-standing model describes the distribution of attentional resources using a “zoom lens” metaphor (Cave & Bichot, 1999; Eriksen & James, 1986) with two defining features. First, attention is spread across space with a gradient of processing quality that declines at locations farther away from the central focus (Downing, 1988; Eriksen & James, 1986; Beck & Ambler, 1973). For example, the Eriksen flanker paradigm (Eriksen & James, 1986) has shown that, as irrelevant distractors are presented closer to the center of the attended region, RTs are slowed. Likewise, target processing improves monotonically as the distance from the center of the attended region declines (Downing, 1988). Second, zoom lens models posit an inverse relationship between processing quality and the size of the attended region, such that spatial cueing benefits increase with smaller cued regions (Castiello & Umiltá, 1990; see also LaBerge, 1983).

Although there is ample behavioral evidence supporting the utility of the zoom lens metaphor, it is still debated whether the behavioral findings reflect change in early stages of visual processing, that is, changes in the quality of visual perception through sensory enhancement

(Luck, Heinze, Mangun, & Hillyard, 1990; Mangun & Hillyard, 1987) or changes in the efficiency of postperceptual processes, such as decision or response selection through restricting decision processes to relevant information (e.g., Eckstein, Shimozaki, & Abbey, 2002; Palmer, 1995; Palmer, Ames, & Lindsey, 1993). One indication that the breadth of attentional orienting influences early perceptual stages of processing comes from studies examining neural activity evoked by attended and unattended stimuli (Itthipuripat, Garcia, Rungratsameetaweemana, Sprague, & Serences, 2014; Müller, Bartelt, Donner, Villringer, & Brandt, 2003). For example, preparatory BOLD activity preceding a target has a greater spatial extent in retinotopically mapped visual areas when participants deployed attention to a broader region of space, suggesting that zoom lens effects cannot be fully explained by changes in postperceptual processing (Müller et al., 2003). Although the prior work makes a good case that changes in the breadth of attention affect activity in visual regions, there is a critical gap in these findings that we aimed to address with the present work. Specifically, although those studies documented a larger number of voxels that passed statistical threshold when attention was broadly directed, this empirical pattern does not entail any change in the “spatial selectivity” of that activity. One alternative account, for example, is that there could have been a baseline shift of activity in the measured visual region, such that responses were elevated across all topographically mapped regions, regardless of their spatial mapping; this kind of baseline shift could increase the number of voxels passing statistical threshold—and thus

¹University of Southampton, ²University of Chicago

the spatial extent of the “activated” region—even if there were no change in the underlying selectivity of the global activity in that visual area. To address this gap in the prior evidence in the present work, we employed an inverted encoding analytic approach that provided a more direct assessment of the spatial selectivity of covert spatial attention. Briefly, this approach estimates activity across a set of spatial “channels” that tile the space of possible target positions, thereby revealing the graded patterns of channel activity that are observed when attention is directed to a specific position in the visual field. This method has been used in past work to precisely track locus and timing of covert spatial attention (Foster, Sutterer, Serences, Vogel, & Awh, 2017). Moreover, by quantifying the selectivity of channel activity centered on the attended positions, we could clearly determine whether changes in the breadth of covert spatial attention influenced the spatial selectivity of the neural activity that indexes covert spatial attention.

We cued participants to direct spatial attention toward either a narrow or broad region of space and used EEG measurements of neural oscillations in the alpha frequency band (8–12 Hz) to track the spatial selectivity of the observers’ attentional focus. This followed a large body of evidence showing that alpha activity provides a precise and temporally resolved index of the locus of covert spatial orienting (Foster, Bsaies, Jaffe, & Awh, 2017; Foster, Sutterer, et al., 2017; Samaha, Sprague, & Postle, 2016; Gould, Rushworth, & Nobre, 2011; Marshall, O’Shea, Jensen, & Bergmann, 2015; Rihs, Michel, & Thut, 2007). Although there is robust evidence that alpha activity tracks the centroid of the attended region, however, past work has not determined whether alpha activity also indexes the “breadth” of the selected region of space. To address this question, we used an inverted encoding model to measure the spatial selectivity of alpha activity while participants were cued to direct spatial attention toward either a narrow or broad region of space within a circular array of possible target positions. Alpha activity has been robustly linked to modulations of sensory activity in retinotopically organized regions (e.g., Worden, Foxe, Wang, & Simpson, 2000). To anticipate our results, behavioral data replicated past observations that motivated the zoom lens metaphor for covert spatial attention. Critically, preparatory activity in the alpha frequency band mirrored these behavioral effects. Spatially specific channel activity peaked at the center of the cued region and dropped in a graded fashion with distance away from that point. Moreover, the slope of this attentional gradient was steeper when observers directed attention narrowly, showing that voluntary adjustments of the attentional zoom lens elicit a flexible tuning of spatially selective neural activity that is thought to gate incoming perceptual information. Thus, these findings complement the past neural studies of the attentional zoom lens by demonstrating that voluntary changes in the breadth of spatial attention evoke changes in the spatial selectivity of the neural signals that track covert

spatial attention. In turn, these findings solidify a perceptual gating model of how adjustment in the zoom lens of attention shapes visual processing.

METHODS

Experiment 1

Participants

Twenty-three volunteers naïve to the objective of the experiment participated for payment (15 USD per hour). One additional participant had to be excluded because they decided to abort the experiment prematurely. Participants were aged 18–31 years ($M = 23.4$, $SD = 3.8$) and reported normal or corrected-to-normal visual acuity as well as normal color vision. Fifteen participants were women, and three were left-handed. The experiment was conducted with the written understanding and consent of each participant.

Stimuli, Apparatus, and Procedure

Participants were seated in a comfortable chair in a dimly lit, electrically shielded and sound-attenuated chamber. Participants put their head in a chinrest at a distance of 75 cm from the screen. They responded with button presses on a standard keyboard that was placed in front of them. Stimulus presentation and response collection were controlled by a Windows PC using PsychToolBox 3 routines in MATLAB (Version 8.6.0). All stimuli were presented on an LCD-TN screen (BenQ XL2430-B).

All stimuli were presented on a gray (RGB: 100-100-100) background. A trial started with a “ready screen” showing a central gray fixation dot (diameter: 0.3° visual angle) slightly brighter than the background (RGB: 128-128-128) (Figure 1). Participants initiated a trial by pressing the space bar, which turned the fixation dot into a fixation cross (0.4° diameter) of the same color. After 500 msec, a cue display appeared for 600 msec. The cue display comprised of a central fixation cross and eight circles (0.2° diameter) arranged on an imaginary circle (2.4° radius) in a way that four circles appeared in the left and four in the right hemifield. Three adjacent circles were blue (0-185-255), and the remaining five circles were green (0-205-0) or vice versa (counterbalanced across participants). Blue and green were determined to be isoluminant, but note that luminance can vary for different monitors/hardware. Importantly, the cues to attend narrow and broad regions were physically identical. Participants were instructed to attend the central one of the three blue/green circles in the cue size 1 (CS1) condition and to attend all three blue/green circles in the cue size 3 (CS3) condition. The cued breadth of focus alternated between blocks of 64 trials. After the cue display, a blank screen, only showing the fixation cross, was shown for 200 msec. Then, a probe display was shown for 50 msec. The probe display showed a central fixation

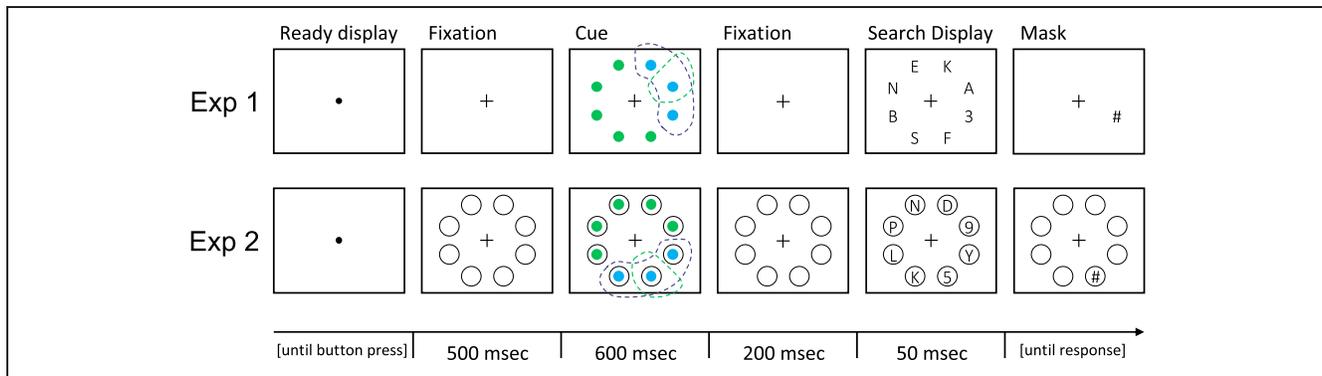


Figure 1. Trial procedure used in Experiment 1 (top) and Experiment 2 (bottom). The green dotted line indicates the cued location in cue size 1 (CS1) blocks, the blue dotted line indicates the cued locations in cue size 3 (CS3) blocks. Dotted lines were not visible to participants. Note that the upper trial shows an example with a valid cue in the “CS3” condition but an invalid cue in the “CS1” condition. The lower example shows a trial that would be valid for both cue sizes. Experiment 2 used placeholders (not present in Experiment 1) throughout an entire trial to facilitate attention deployment at clearly defined spatial locations.

cross, seven letters (randomly picked without repetition from these letters: A, C, D, E, F, H, K, L, M, N, P, R, S, T, U, V, W, X, Y, Z) and one digit (1–9) in the same positions as the circles in the cue display. The probe letters/digit had a shade of gray that was determined in a staircasing procedure (see below), ranging between 102-102-102 and 250-250-250. The digit appeared at the cued location in 75% of the trials and at any other location in the remaining 25% of the trials (75% cue validity). The probe display was followed by a mask display showing a central fixation cross and a pound sign (#) at the location where the digit was shown. The mask display stayed on the screen until participants reported the identity of the digit; participants were instructed to report the digit as accurately as possible (unsped) and to ignore all letters. They entered the digit by pressing one of the numpad keys on a standard Windows keyboard. After an intertrial interval of 500 msec, a new “ready screen” with a central fixation dot indicated that participants could start the next trial. Participants completed 20 blocks of 64 trials each (total of 1280 trials). Feedback about their performance (percent correct) was provided to participants after each block.

Control for Gaze Position

Gaze position was tracked at a sampling rate of 1000 Hz for both eyes with an EyeLink 1000+ eye tracker (SR Research Ltd.). A direct gaze feedback violation procedure was applied from 450 msec after the trial start (fixation cross onset) until the onset of the mask display, that is, for 900 msec. If a participant’s gaze was not within 1.5° of the center of the fixation cross during that time or if they blinked, the trial was aborted, and a message “eye movement” (or “blink”) was presented on the screen before a ready screen indicated the restart of the trial. The remaining trials were shuffled so as to put the aborted trial in a random position within the sequence and make

its reappearance unpredictable. Any detected gaze violation extended the experiment by one trial.

Staircase Procedure

All participants underwent a staircase procedure in a separate session, typically 1–3 days before the main experiment, to determine the contrast between the background and the probe items that would ensure a performance neither at the floor nor at the ceiling level. The task was identical to the main task described above with the difference that only one circle in the cue display had a deviating color and participants were instructed to always attend that location. In the staircase procedure, the cue had 100% validity. Participants started with a contrast of 150 (difference in RGB values between background and probe items). Whenever participants responded correctly, contrast was reduced by 10% or at least two RGB values. Whenever participants responded incorrectly, contrast was increased by 20% or at least three RGB values. This was done for 128 trials, and the ideal contrast was determined as the average contrast of the last 11 trials of the staircase procedure. The staircase procedure set participants at a contrast level that would correspond to an average performance of approximately two out of three correct (chance level = $1/9 = 11\%$). On the same day as the main task, the staircase procedure was repeated for 32 trials, starting with the contrast determined in the initial session. During the main task, contrast was fixed.

EEG Recording

EEG was recorded with Ag–AgCl active electrodes (BrainProducts actiCap) from 32 scalp sites (according to the International 10/20 System: FP1/2, F7/8, F3/4, Fz, FC5/6, FC1/2, C3/4, Cz, TP9/10, CP5/6, CP1/2, P7/8, P3/4, PO7/8, PO3/4, Pz, O1/2, Oz). Horizontal and vertical

EOGs were recorded with passive electrodes bipolarly at ~ 1 cm from the outer canthi of the eyes and from above and below the observers' right eye, respectively. Fpz served as the ground electrode, and all electrodes were referenced to TP10 and rereferenced offline to the average of all electrodes. Impedances for active electrodes were kept below 10 k Ω . Sampling rate was 1000 Hz, with a high cutoff filter of 250 Hz and a low cutoff filter of 0.01 Hz (half power cutoff, 12 dB roll-off).

Data Analysis

Behavioral data. Accuracy was analyzed as a function of cue size (1 vs. 3) and as a function of distance of the target digit to the cued location (CS1) or the center of the cue (CS3). Distances were 0, +1, +2, +3, -1, -2, -3, ± 4 . A distance of 0 (CS1) or a distance of 0 or 1 (CS3) refers to valid trials. This led to 16 accuracy values for each participant that were forwarded to a 2×8 ANOVA with the repeated-measures factors Cue size and Distance. Furthermore accuracy as a function of cue–target distance was fit to a sine function separately for CS1 and CS3.

EEG data. EEG was segmented offline over a 1600-msec epoch, including a 400-msec prestimulus baseline with epochs time-locked to cue display onset. Trials with both correct and incorrect responses were used. Trials with eye-related artifacts from -200 to 800 msec were excluded from the analysis (Experiment 1: 2.8%, $SD = 3.0\%$; Experiment 2: 5.7%, $SD = 4.8\%$). To identify eye-related artifacts, eye-tracking data were first baselined identically to EEG data (i.e., subtraction of the mean amplitude of x and y coordinates for the time from -200 to 0 msec). Then, the Euclidian distance from the fixation cross was calculated from baselined data. We identified saccades with a step criterion of 0.6° (comparing the mean position in the first half of a 50-msec window with the mean position in the second half of a 50-msec window; window moved in 20-msec steps). We identified drifts by eye-tracking data indicating a distance from the fixation of $> 1^\circ$. Both eyes had to indicate an eye-related artifact for a trial to be excluded from analysis. Three participants in Experiment 1 and two participants in Experiment 2 did not have eye-tracking data available. For these participants, EOG was used instead (100 μV absolute voltage difference from baseline or 40 μV step criterion, same window technique as for eye-tracking data described above). In addition, we rejected trials in which any EEG channel showed a voltage of more than 100 μV or less than -100 μV . Any electrode showing more than 50 such trials was rejected from the analysis, and the individual electrode rejection was run again disregarding that electrode.

To isolate alpha-specific activity, raw EEG segments were band-pass filtered (8–12 Hz) using a two-way least-squares finite impulse response filter (*eegfilt.m* from

EEGLAB Toolbox) and then Hilbert-transformed. To reconstruct spatially selective channel tuning functions (CTFs) from the topographic distribution of oscillatory power across electrodes, we used an inverted encoding model (IEM; Foster, Sutterer, et al., 2017). The IEM assumes that power measured at each electrode reflects the weighted sum of eight spatial channels (representing neuronal populations), each tuned for one of the eight target positions. In a training stage, two thirds (equaled for the 2×8 combinations of cue size and position of the cue/cue center) of the segmented, filtered trials were used to estimate the weights in a least-squares estimation. In the test stage, the model was inverted to transform the remaining third of trials (again, equaled for the 2×8 trial types) into estimated channel responses using the previously determined weights. This means that a common set of training data (sampled equally from the narrow and the broad conditions) was used to estimate the channel responses in the narrow and broad conditions separately. This procedure precludes the possibility that any observed effects of attentional breadth are due to differences that arose during training or because of using distinct “basis sets” for the two conditions. The assignment of trials to training/test stage was done for 1000 iterations and the resulting channel-response profiles were averaged across iterations to achieve a better signal-to-noise ratio (Foster, Sutterer, et al., 2017). The eight channel response functions were shifted to a common center and averaged to obtain the CTF. To compare the channel responses between conditions, the CTF was averaged across an early epoch of 200–400 msec and a late epoch of 600–800 msec. The early time window was chosen to track attention processes that are late enough to not reflect the response to the physical onset of the cues but early enough to reflect early attention deployment. The N2pc component as an ERP measure of attention, for example, typically falls into this window. The late time window was chosen to track sustained attention at the moment just before the target appeared. The data were forwarded to a 2×2 ANOVA with the repeated-measures factors Cue size (1 vs. 3), and the Distance of the cued location to the location a channel was optimally tuned to (0 through ± 4), analogously to the accuracy analysis described above. Additionally, similarly to accuracy (see above), the CTF was fit to a sine function to obtain amplitude, dispersion, and baseline parameters. In addition, the slope of the CTFs (estimated by linear regression computed for each time point) was used as a metric to compare the selectivity between CS1 and CS3; the higher the slope, the greater the spatial selectivity. The slope was averaged for eight epochs (100 msec time windows from 0 to 800 msec) separately for CS1 and CS3 and tested against zero and compared between CS1 and CS3.

Simulating CTF slopes for CS3. Although a shallower CTF slope in the CS3 condition can be interpreted as

reflecting a broader attended region, we also considered the possibility that participants may show an equally narrow focus of attention in both conditions but deployed that focus probabilistically to one of the three locations in the CS3 condition. Thus, deploying a narrow attentional focus across the three positions might be able to mimic the effects of a single broader focus of attention. To test this possibility, we examined whether the CTF slopes in the CS3 condition could be recreated by probabilistic switching of a narrow focus of attention across multiple positions. For instance, if observers narrowly attended each cued position one third of the time in the CS3 condition, this might mimic the gradient produced by a single broader focus of attention. To test this account, we first determined how often the CS1 CTF profile should be directed to each of the three positions in the CS3 condition to obtain the best possible match with the observed CTF in the CS3 condition. We used the data from the CS1 condition and calculated the weighted sum of channel activity with a varying ratio of how often the central location within the cued area would be attended by a “switching participant.” We varied the ratio from one out of three (equally likely attending the central

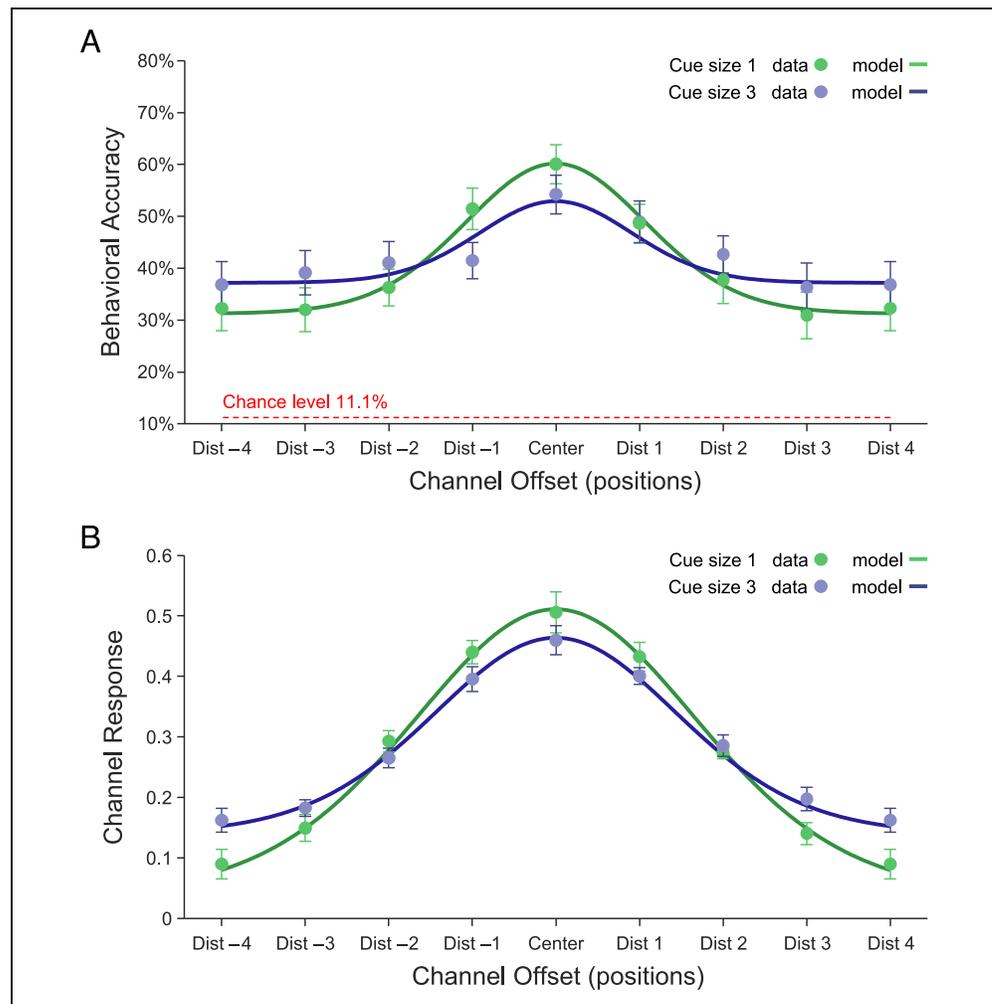
as the lateral locations) to 100% (never attending the lateral locations, i.e., identical to the CS1 condition). The ratio for which the sum of the squared difference between the observed and simulated CS3 condition for each channel offset (measure of how dissimilar the CTFs were) was minimal was used to calculate the slope for the simulated cue size 3 condition.

Results Experiment 1

Behavioral Results

Accuracy was reliably above chance level (11.1%) for all 16 combinations of cue size and distance to the (center of the) cued location (all p s < .001). Average accuracy did not differ reliably across cue sizes ($M_{CS1} = 41.3\%$ vs. $M_{CS3} = 42.7\%$), $F(1, 22) = 3.6$, $p = .071$, $\eta^2 = .141$. Accuracy varied as a function of the target distance to the cued location, $F(7, 154) = 30.1$, $p < .001$, $\eta^2 = .577$ (see Figure 2). Accuracy was highest for the central cued location and dropped to more distant locations ($M_{dist0} = 57.2\%$ vs. $M_{dist4} = 34.7\%$; see Figure 2). How accuracy was affected by the distance to the cued location varied

Figure 2. Experiment 1. (A) Mean behavioral accuracy for cue size 1 (green dots) and cue size 3 (blue dots) blocks. (B) Estimated channel responses for CS1 (green dots) and CS3 (blue dots) blocks for an epoch from 600 to 800 msec. Error bars show *SEMs*, corrected for individual differences (Cousineau, 2005). Solid lines show a sine function fitted to the data.



as a function of cue size, yielding an interaction of Distance and Cue size, $F(7, 154) = 6.5, p < .001, \eta^2 = .229$. Within-participant contrasts revealed that the interactive pattern followed a quadratic trend, $F(1, 22) = 14.5, p = .001, \eta^2 = .398$, but not a linear ($p = .594$) or cubic ($p = .379$) trend. Follow-up analyses showed that accuracy varied as a function of distance for both CS1, $F(7, 154) = 28.5, p < .001, \eta^2 = .565$, and CS3, $F(7, 154) = 11.6, p < .001, \eta^2 = .345$. The slope of the accuracy-distance function was steeper for CS1 (7.4%) than for CS3 (4.3%), $t(22) = 4.9, p < .001$.

Channel Tuning Functions

First epoch (200–400 msec). Channel activity was reliably above zero for all 16 combinations of cue size and distance to the cued location (all $ps < .001$). Average channel responses were not affected by cue size ($M_{CS1} = 0.29$ vs. $M_{CS3} = 0.29$), $F(1, 22) = 1.2, p = .288, \eta^2 = .051$. Channel responses varied as a function of the target distance to the cued location, $F(7, 154) = 11.9, p = .001, \eta^2 = .351$. The highest channel response was found for the cued location and dropped to more distant locations ($M_{\text{dist}0} = 0.36$ vs. $M_{\text{dist}4} = 0.23$; see Figure 3 for a CTF as a function of time). There was no interaction of Distance and Cue size, $F(7, 154) = 0.6, p = .627, \eta^2 = .025$.

Second epoch (600–800 msec). Channel activity was reliably above zero for all 16 combinations of cue size and distance to the cued location (all $ps < .001$). Average channel responses were not affected by cue size ($M_{CS1} = 0.29$ vs. $M_{CS3} = 0.29$), $F(1, 22) = 0.6, p = .451, \eta^2 = .026$. Channel responses varied as a function of the target distance to the cued location, $F(7, 154) = 58.5, p < .001, \eta^2 = .727$. The highest channel response was found for the cued location and dropped to more distant locations ($M_{\text{dist}0} = 0.48$ vs. $M_{\text{dist}4} = 0.14$; see Figure 3). An interaction of Distance and Cue size, $F(7, 154) = 3.3, p = .032, \eta^2 = .130$, showed that how channel responses were affected by distance to the cued location varied as a function of cue size. Within-participant contrasts revealed that the interactive pattern followed a quadratic trend, $F(1, 22) = 7.2, p = .014, \eta^2 = .246$, but not a linear ($p = .656$) or cubic ($p = .993$) trend. Follow-up analyses showed that channel response varied as a function of distance for both CS1, $F(7, 154) = 41.5, p < .001, \eta^2 = .653$, and CS3, $F(7, 154) = 37.3, p < .001, \eta^2 = .629$.

CTF Slopes

No reliable slopes were found for CS1 of 0–100 and 100–200 msec (all $ps \geq .244$), but for all succeeding time windows (all $ps \leq .048$). For CS3, no time window before 300 msec showed a reliable slope (all $ps \geq .078$), but all time windows from 300 msec on showed a reliable slope (all $ps \leq .003$). Differences in the time slope between CS1

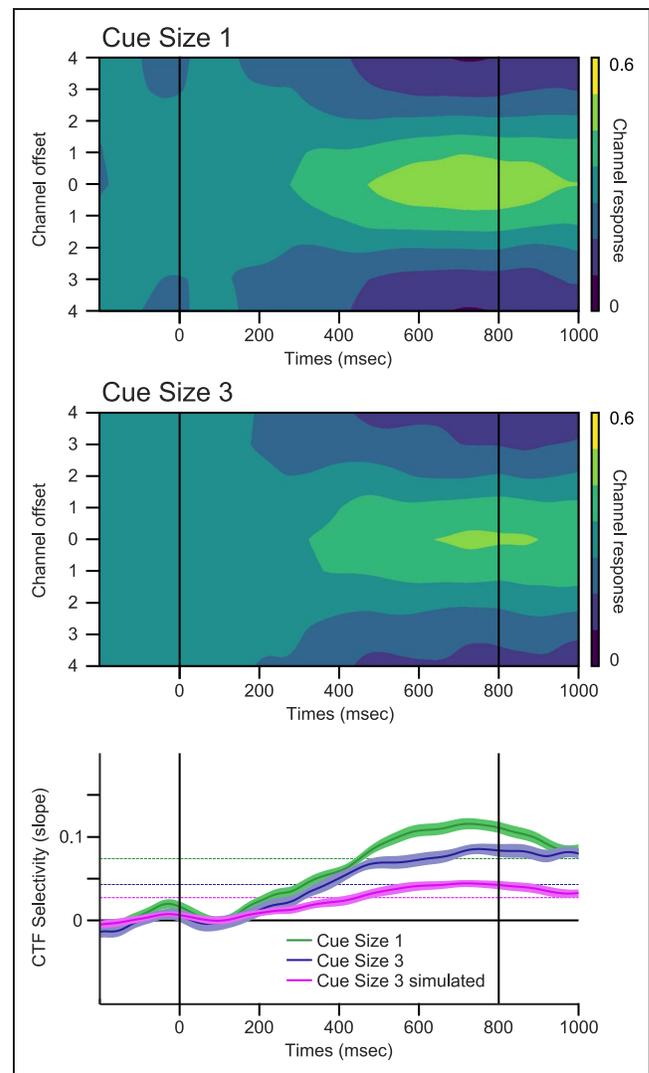


Figure 3. CTFs in Experiment 1 for cue size 1 (top) and cue size 3 (middle). The slope of these CTFs provides a measure of the spatial selectivity of alpha activity that indexes the attended region (bottom). A narrower focus of attention (green lines) induced steeper slopes than a broader focus (blue lines). The pink line represents the simulated slope for the cue size 3 condition under the assumption that, rather than applying a broad spotlight of attention encompassing all three locations, a narrow spotlight of attention encompassing one location only is directed toward one of the three locations on a trial-by-trial basis. The dotted lines represent the slope of the behavioral functions. Shaded envelopes represent the SEM corrected for individual differences (Cousineau, 2005).

and CS3 were not observed before 600 msec (all $ps \geq .060$), but for all succeeding time windows (all $ps \leq .044$).

The fit between the simulated and observed CS3 was best under the assumption that attention was focused on each of the three positions equally often (i.e., ratio = 1/3). The simulated CS3 slope was not reliable before 300 msec (all $ps \geq .052$), but for all succeeding time windows (all $ps \leq .002$). CS3 and simulated CS3 did not show reliable differences before 300 msec (all $ps \geq .274$), but CS3 had a steeper slope than simulated CS3 for all

succeeding time windows (all $ps \leq .028$). This suggests that probabilistic switching of the CTF from the CS1 condition could not match the slope of the observed CTF in the CS3 condition. We ran an analogous analysis for the accuracy data. As for the CTF data, the optimal ratio for the central location was one out of three and the slope was reliably different between the observed (4.3%) and simulated (2.8%) CS3 condition ($p = .024$). Thus, probabilistic switching of a narrow attentional focus does not provide an adequate explanation of the broader tuning observed in both behavioral and neural data when observers were cued to attend a wider region of space.

The behavioral and EEG findings followed a similar pattern. In the CS1 condition, target discrimination accuracy was higher in the center of the cued region than in the CS3 condition, and accuracy dropped more quickly with increasing distance from the center of the cued region in the CS1 condition than in the CS3 condition. Likewise, from 500 msec on, the slope of the CTF was steeper in the CS1 condition than in the CS3 condition, suggesting a sharper drop-off in the gradient of attention around the cued region. Interestingly, reliable slopes for each cue size were observed before a differential slope was found, namely, from 200 msec on (CS1) or 300 msec on (CS3). This falls in line with past estimates of the time course to deploy attention following symbolic central cues (Feldmann-Wüstefeld & Schubö, 2013; Müller & Rabbitt, 1989). Apparently participants shifted their attention to the peripheral location before they then adjusted the size of the attended region.

Experiment 2

Experiment 2 served two purposes. First, we sought to replicate our observation that the profile of spatial channel activity tracks the breadth of covert attentional orienting. Second, we hypothesized that observers might be able to control the breadth of attention more easily if there were physical landmarks in the cued positions during the time between cue and target. To that end, we presented ring-shaped placeholders at all eight positions throughout the cue–target interval.

Participants

Twenty-seven volunteers naïve to the objective of the experiment participated for payment (15 USD per hour). Participants were aged 19–37 years ($M = 23.9$, $SD = 4.9$) and reported normal or corrected-to-normal visual acuity as well as normal color vision. Fifteen participants were women, and one was left-handed. The experiment was conducted with the written understanding and consent of each participant.

Stimuli, Apparatus, and Procedure

Experiment 2 was identical to Experiment 1 except for these differences: Throughout the entire trial (i.e., except

for the intertrial interval and the “ready display”), eight empty gray circles (1.7°) with the same luminance as the fixation cross served as placeholders. This was intended to help observers to attentionally lock onto cued locations and reduce spatial uncertainty. Furthermore, the online eye-tracking procedure allowed observers to only deviate 1.2° from fixation but allowed any deviation from fixation for a maximum of 50 msec (this was done to optimize the feedback procedure and avoid false alarms, i.e., signaling eye movements to participants when in fact there were none). This allowed for noise in the eye-tracking signal and provided more accurate feedback for detecting eye movements. Note that the post hoc artifact rejection was identical for Experiments 1 and 2, ensuring a similar data quality. Cue validity was 90%, to further increase participants’ motivation to use the cues. In the staircasing procedure, the average of the last 21 contrast values was used in the main experiment to get a more reliable estimate of an appropriate difficulty level for a given participant.

Results Experiment 2

Behavioral Results

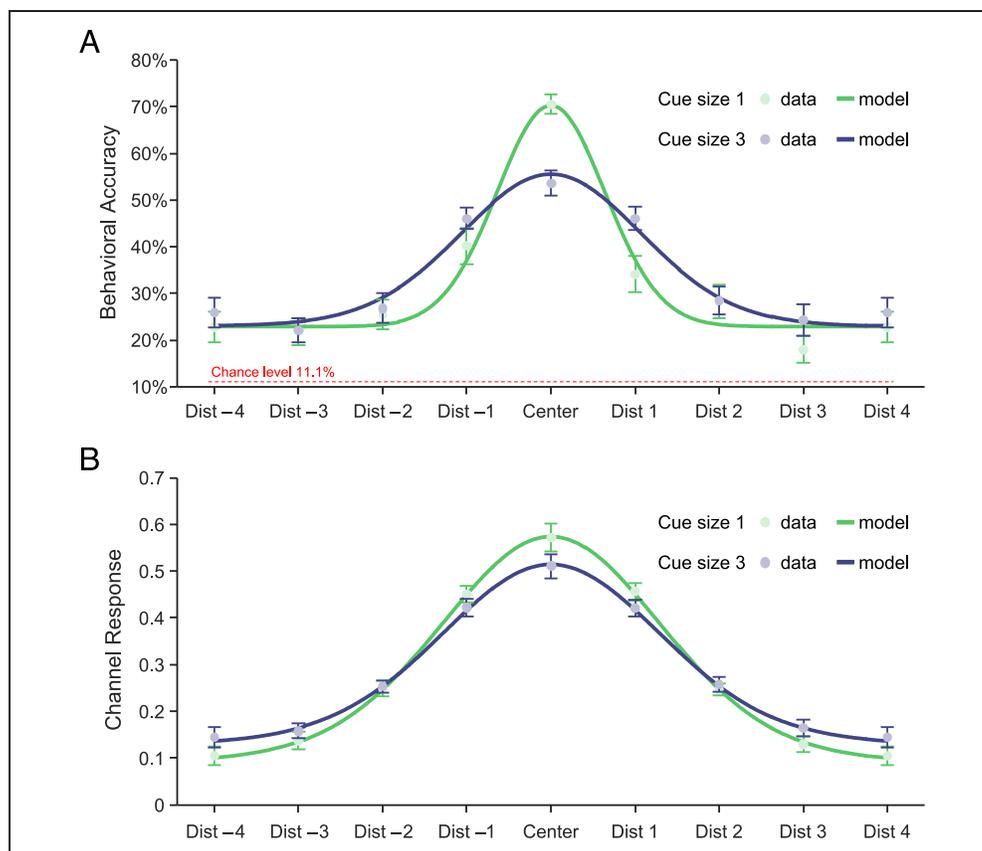
Accuracy was reliably above chance level (11.1%) for all 16 combinations of cue size and distance to the (center of the) cued location (all $p_{1t} < .001$). There was no main effect of cue size ($M_{CS1} = 32.8\%$ vs. $M_{CS3} = 34.2\%$), $F(1, 22) = 2.4$, $p = .135$, $\eta^2 = .084$. Accuracy varied as a function of the target distance to the cued location, $F(7, 182) = 67.5$, $p < .001$, $\eta^2 = .722$. Accuracy was highest for the cued location and dropped to more distant locations ($M_{dist0} = 62.1\%$ vs. $M_{dist4} = 24.5\%$; see Figure 4). The effect of distance from the cued location varied as a function of cue size, yielding an interaction of Distance and Cue size, $F(7, 182) = 6.4$, $p < .001$, $\eta^2 = .198$. Within-participant contrasts revealed that the interactive pattern followed a marginally quadratic trend, $F(1, 22) = 3.4$, $p = .078$, $\eta^2 = .115$, but not a linear ($p = .408$) or cubic ($p = .798$) trend. Follow-up analyses showed that accuracy varied as a function of distance for both CS1, $F(7, 182) = 39.0$, $p < .001$, $\eta^2 = .600$, and CS3, $F(7, 182) = 36.9$, $p < .001$, $\eta^2 = .587$. The slope of the accuracy–distance function was steeper for CS1 (11.2%) than for CS3 (7.6%), $t(26) = 3.4$, $p = .002$.

Channel Tuning Functions

First epoch (200–400 msec). Channel activity was reliably above zero for all 16 combinations of cue size and distance to the cued location (all $ps < .001$). Average channel responses were not affected by cue size ($M_{CS1} = 0.29$ vs. $M_{CS3} = 0.29$), $F(1, 26) = 0.3$, $p = .614$, $\eta^2 = .010$. Channel responses varied as a function of the distance from the center of the cued location, $F(7, 182) = 34.8$, $p < .001$, $\eta^2 = .573$. The highest channel response was found for the cued location and dropped monotonically as the

Figure 4. Experiment 2.

(A) Mean behavioral accuracy for cue size 1 (CS1) (green dots) and cue size 3 (CS3) (blue dots) blocks. (B) Estimated channel responses for CS1 (green dots) and CS3 (blue dots) blocks for an epoch from 600 to 800 msec. Error bars show standard errors of the mean, corrected for individual differences (Cousineau, 2005). Solid lines show a sine function fitted to the data.



distance from the cued position increased ($M_{\text{dist}0} = 0.41$ vs. $M_{\text{dist}4} = 0.18$; see Figure 5 for a CTF as a function of time). An interaction of Distance and Cue size, $F(7, 182) = 4.5$, $p = .012$, $\eta^2 = .148$, showed that the effect of distance on channel responses depended on cue size. Within-participant contrast revealed that the interactive pattern followed a quadratic trend, $F(1, 26) = 13.9$, $p = .002$, $\eta^2 = .348$, but not a linear ($p = .805$) or cubic ($p = .198$) trend. Follow-up analyses showed that channel response varied as a function of distance for both CS1, $F(7, 182) = 37.5$, $p < .001$, $\eta^2 = .590$, and CS3, $F(7, 182) = 17.9$, $p < .001$, $\eta^2 = .407$.

Second epoch (600–800 msec). Channel activity was reliably above zero for all 16 combinations of cue size and distance to the cued location (all $ps < .001$). Channel responses were not affected by cue size ($M_{\text{CS1}} = 0.29$ vs. $M_{\text{CS3}} = 0.29$), $F(1, 26) < 0.1$, $p = .986$, $\eta^2 < .001$. Channel responses varied as a function of the target distance to the cued location, $F(7, 182) = 89.5$, $p < .001$, $\eta^2 = .776$. The highest channel response was found for the cued location and dropped to more distant locations ($M_{\text{dist}0} = 0.54$ vs. $M_{\text{dist}4} = 0.13$; see Figure 5). An interaction of Distance and Cue size, $F(7, 182) = 4.6$, $p = .007$, $\eta^2 = .151$, showed that how channel responses were affected by distance to the cued location varied as a function of cue size. Within-participant contrast revealed

that the interactive pattern followed a quadratic trend, $F(1, 22) = 10.6$, $p = .003$, $\eta^2 = .289$, but not a linear ($p = .840$) or cubic ($p = .322$) trend. Follow-up analyses showed that channel response varied as a function of distance for both CS1, $F(7, 182) = 78.5$, $p < .001$, $\eta^2 = .751$, and CS3, $F(7, 182) = 58.5$, $p < .001$, $\eta^2 = .692$.

CTF Slopes

No reliable slopes were found for CS1 from 0 to 100 ($p = .461$), but for all succeeding time windows (all $ps \leq .010$). For CS3, no time window before 200 msec showed a reliable slope (all $ps \geq .383$), but all time windows from 200 msec on showed a reliable slope (all $ps \leq .023$). Differences in the time slope between CS1 and CS3 were not observed before 200 msec (all $ps \geq .190$) and from 300–500 msec (all $ps \geq .116$), but for all other time windows (all $ps \leq .025$). The CTF similarity for the simulated and observed CS3 was highest under the assumption that attention was focused on the central location in 42% of the trials and on each of the lateral location in 29% of the trials. For the simulated CS3 slope, no reliable slopes were found from 0 to 100 msec ($p = .377$), but for all succeeding time windows (all $ps \leq .008$). Differences in the time slope between CS3 and simulated CS3 were not observed before 300 msec (all $ps \geq .829$), but for all succeeding time windows (all $ps \leq .001$). This replicated

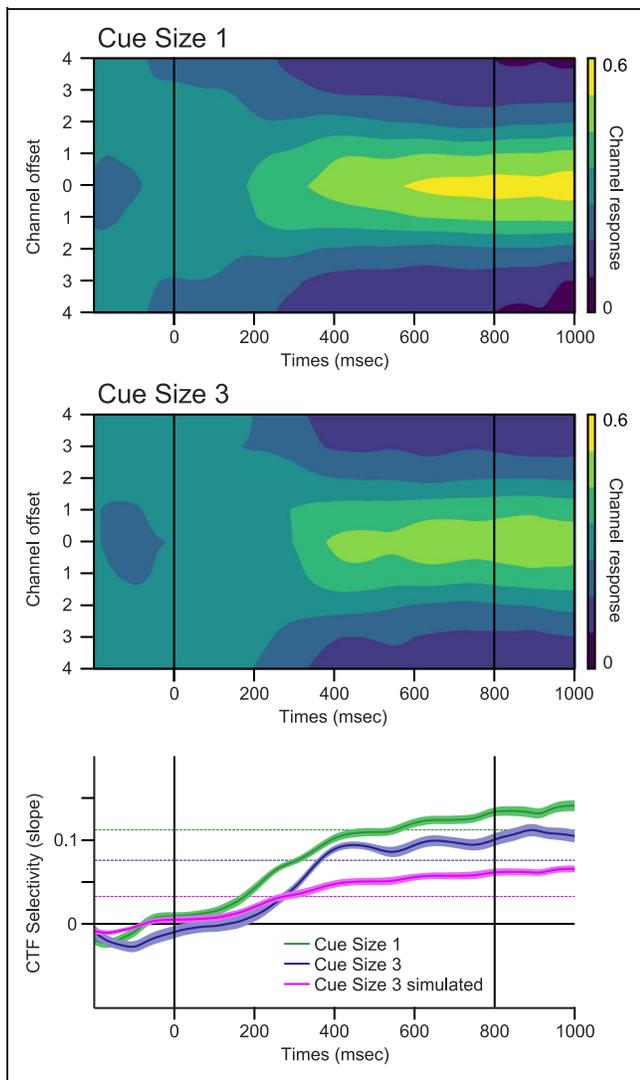


Figure 5. CTFs in Experiment 2 for CS1 (top) and CS3 (middle). The slope of these CTFs is used as a measure of how selectively a channel responds to cueing of the location it represents (bottom). Smaller cues (green lines) induce steeper slopes than larger cues (blue lines). The pink line represents the simulated slope for the CS3 condition under the assumption that rather than applying a broad spotlight of attention encompassing all three locations, a narrow spotlight of attention encompassing one location only is directed toward one of the three locations on a trial-by-trial basis. The dotted lines represent the slope of the behavioral functions. Shaded envelopes represent the *SEM* corrected for individual differences (Cousineau, 2005).

findings from Experiment 1 and suggests that probabilistic switching of the CTF from the CS1 condition could not match the slope of the observed CTF in the CS3 condition. Again, we carried out the same analysis for the behavioral data. The optimal ratio for the central location was 45%, and the slope was reliably different between the observed (7.6%) and simulated (4.6%) CS3 condition ($p = .002$). Thus, the data from the CS3 condition could not be reproduced by probabilistic switching of the narrow focus observed in the CS1 condition. In summary, both

Experiments 1 and 2 show that participants oriented more broadly in the CS3 condition and that spatially specific activity in the alpha frequency band tracked these changes in the size of the zoom lens of attention.

It is noteworthy that reliable effects of cue size at an earlier point in time in Experiment 2 (200 msec) than in Experiment 1 (600 msec). This raised the possibility that landmarks facilitated more efficient control over the breadth of orienting. To test this apparent difference between Experiments 1 and 2 more directly, we ran an exploratory analysis to compare the slope differences. Using the same 100-msec intervals, we compared the differential slope (CS1 minus CS3) between experiments with a *t* test for independent samples. There was no reliable difference for any time interval (all $ps \geq .056$). Thus, we did not find robust evidence that the effect of cue size emerged earlier in Experiment 2 than in Experiment 1.

GENERAL DISCUSSION

There has been longstanding behavioral evidence suggesting that observers can exert voluntary control over the size of attention's "zoom lens," such that limited resources for visual selection are spread over a narrow or a broad region of space (Cave & Bichot, 1999; Eriksen & James, 1986). In this study, participants were cued to direct spatial attention toward either a narrow or broad region of space within a circular array of possible target positions. In line with the zoom lens model (Castiello & Umiltá, 1990; Eriksen & James, 1986; LaBerge, 1983), a gradient of accuracy was found with best performance at the cued location and a drop-off toward more distant locations. Second, the gradient varied as a function of cue size, such that a narrow attentional cue elicited a faster drop in discrimination accuracy as the distance from the center of the cued region increased.

Using an EEG measure of the breadth of covert spatial attention, the present work extends previous evidence for the zoom lens model in multiple ways. Although past studies had documented changes in the spatial extent of activity in visual cortex following adjustments to the zoom lens of attention (Itthipuripat et al., 2014; Müller et al., 2003), this empirical pattern could not distinguish between baseline shifts in visual activity and changes in the spatial selectivity of that activity. To discriminate between these alternatives, we used an inverted encoding analytic approach that enabled a more direct measurement of the spatial selectivity of attention-based neural activity. Specifically, we focused on preparatory changes in neural oscillations in the alpha frequency band, a brain rhythm that has been robustly linked with modulations of incoming sensory information. The inverted encoding analysis showed that spatially selective alpha activity exhibited both defining features of the zoom lens model (LaBerge, 1983). First, channel activity peaked at the center of the cued region and dropped in a graded fashion with distance away from that point, suggesting that

attention is spread across space with a gradient of processing quality that declines at locations farther away from the central focus. Second, the slope of this attentional gradient was steeper when observers directed attention narrowly, suggesting that the size of the attentional gradient can be flexibly adjusted and processing quality declines as the attended region grows broader.

Furthermore, this study advances our knowledge of how neural oscillations in the alpha frequency band are related to the control of covert attention. Although prior work had clearly established that alpha activity enables precise tracking of the “location” of the attended region (Foster & Awh, 2019; Foster, Sutterer, et al., 2017; Foxe & Snyder, 2011; Jensen & Mazaheri, 2010; Worden et al., 2000), our application of an inverted encoding analysis provided new evidence that this neural signal indexes changes in the “breadth” of the attended region.

Although there is a large body of literature that documents behavioral effects that are in line with a zoom lens model, such findings can be explained in two distinct ways. They could reflect changes in the quality of visual perception or changes in the efficiency of postperceptual processes, such as decision or response selection. For example, “decision noise models” assume that attention deployment enhances performance by restricting decision processes to relevant information without affecting perceptual processing (Eckstein et al., 2002; Palmer, 1995; Palmer et al., 1993). At the same time, there is clear evidence that spatial attention can also influence earlier visual stages of processing. For example, Mangun and Hillyard (1987) used EEG recordings to show that early visually evoked potentials that reflect sensory processing (Luck et al., 1990) were largest at cued locations and monotonically decreased at farther away location, suggesting that spatial attention modulates the flow of visual sensory information. However, these results only show neural evidence for a gradient of attention at the time of target onset. By contrast, our study provides clear EEG evidence for an “anticipatory” gradient of attention. Because of the high time resolution of EEG measures, we could track the gradient over time, and because IEMs can track attention in the absence of stimulation, we could track the gradient before target onset. A shift of attention to the cued region was observed relatively early, after around 200 msec (or after 100 msec in the CS1 condition of Experiment 2). This presumably reflected an exogenous shift of attention toward the peripheral cue, three dots of unique color. However, it took up to 600 msec for the gradients induced by small and large cues to diverge. Given that the cues for narrow and broad focus conditions were physically identical, our data suggest that it took several hundred milliseconds for observers to exert voluntary control over the breadth of the attended region, a latency that dovetails with past studies of the latency with which observers can orient in response to symbolic cues (Feldmann-Wüstefeld & Schubö, 2013; Müller & Rabbitt, 1989). Although the time course is in

line with past measures of endogenous orienting “toward a location” in the visual field, it is interesting to find a similar time course for adjusting the breadth of attention. In fact there is an interesting parallel in the time course of the emergence of attentional gradients: When individuals are endogenously cued to a location in a homogeneous texture, they attend the entire texture after 200 msec before they can focus their attention on the actual location from 400 msec on (Feldmann-Wüstefeld & Schubö, 2013). Similarly to this bottom-up induced (through texture) gradient of attention, our neural data suggest a similar time course of a top-down induced (through spatial expectancy) gradient of attention.

Because the spatial selectivity of alpha-band activity tracked the gradient and size of the attended region before target onset, our data suggest that the behavioral zoom lens effects reported in the literature reflect a difference in the preparatory stance of the attentional system that cannot be explained by changes in decision efficiency or other postperceptual factors. Thus, our data are the first observation of neural evidence for both properties of the zoom lens model: a gradient of processing quality that declines at locations farther away from the central focus and an inverse relationship between processing quality and the size of the attended region (Castiello & Umiltá, 1990; LaBerge, 1983). Importantly, we also went beyond previous studies by examining whether the neural response pattern that coincided with a broader attended region could be explained by a probabilistic shift of attention to only one of the locations within the region and analysis that exploited the more direct measure of spatial selectivity provided by the inverted encoding analytic approach. We found that the more shallow slope of alpha-band CTF functions could not be explained by any distribution of foci of attention of similar size as in the narrow condition. Rather, our results suggest that the shallow slope is indeed indicative of a broader attended region when large cues, providing less spatial certainty about an upcoming target locations, are presented. Is it possible that we observed lower channel activity in the CS3 condition because of more broadband noise in that condition? This is unlikely because, in a simulation by Sutterer, Foster, Adam, Vogel, and Awh (2019), it was tested whether greater noise could explain channel tuning selectivity differences (using a very similar inverted encoding approach as in this study) for different memory loads in a change detection task. They added Gaussian noise to simulated data in one condition and found that mean CTF selectivity was not affected.

To conclude, an inverted encoding analysis of alpha activity provided direct evidence that the spatial selectivity of this attention-based neural signal responds to voluntary changes in the breadth of covert spatial attention. These findings enrich our understanding of the links between alpha activity and covert spatial attention, while also providing important new constraints for cognitive models of the attentional zoom lens.

Acknowledgments

This work was supported by National Institute of Mental Health grant 2R01MH087214-06A1.

Reprint requests should be sent to Tobias Feldmann-Wüstefeld, Highfield Campus, School of Psychology, Southampton, United Kingdom, or via e-mail: tobias.fw@soton.ac.uk.

REFERENCES

- Beck, J., & Ambler, B. (1973). The effects of concentrated and distributed attention on peripheral acuity. *Perception & Psychophysics*, *14*, 225–230.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.
- Castiello, U., & Umiltá, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica*, *73*, 195–209.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, *6*, 204–223.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*, 42–45.
- Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 188–202.
- Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, *2*, 25–45.
- Eriksen, C. W., & James, J. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*, 225–240.
- Feldmann-Wüstefeld, T., & Schubö, A. (2013). Textures shape the attentional focus: Evidence from exogenous and endogenous cueing. *Attention, Perception, & Psychophysics*, *75*, 1644–1666.
- Foster, J. J., & Awh, E. (2019). The role of alpha oscillations in spatial attention: Limited evidence for a suppression account. *Current Opinion in Psychology*, *29*, 34–40.
- Foster, J. J., Bsaies, E. M., Jaffe, R. J., & Awh, E. (2017). Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Current Biology*, *27*, 3216.e6–3223.e6.
- Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2017). Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention. *Psychological Science*, *28*, 929–941.
- Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*, 154.
- Gould, I. C., Rushworth, M. F., & Nobre, A. C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *Journal of Neurophysiology*, *105*, 1318–1326.
- Itthipuripat, S., Garcia, J. O., Rungratsameetaweemana, N., Sprague, T. C., & Serences, J. T. (2014). Changing the spatial scope of attention alters patterns of neural gain in human cortex. *Journal of Neuroscience*, *34*, 112–123.
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186.
- LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 371–379.
- Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. I. Evidence for early selection. *Electroencephalography and Clinical Neurophysiology*, *75*, 511–527.
- Mangun, G. R., & Hillyard, S. A. (1987). The spatial allocation of visual attention as indexed by event-related brain potentials. *Human Factors*, *29*, 195–211.
- Marshall, T. R., O'Shea, J., Jensen, O., & Bergmann, T. O. (2015). Frontal eye fields control attentional modulation of alpha and gamma oscillations in contralateral occipitoparietal cortex. *Journal of Neuroscience*, *35*, 1638–1647.
- Müller, N., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the “zoom lens” of visual attention. *Journal of Neuroscience*, *23*, 3561–3565.
- Müller, H. J., & Rabbitt, P. M. (1989). Spatial cueing and the relation between the accuracy of “where” and “what” decisions in visual search. *The Quarterly Journal of Experimental Psychology: Section A: Human Experimental Psychology*, *41*, 747–773.
- Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, *4*, 118–123.
- Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 108–130.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, *109*, 160–174.
- Rihs, T. A., Michel, C. M., & Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by α -band EEG synchronization. *European Journal of Neuroscience*, *25*, 603–610.
- Samaha, J., Sprague, T. C., & Postle, B. R. (2016). Decoding and reconstructing the focus of spatial attention from the topography of alpha-band oscillations. *Journal of Cognitive Neuroscience*, *28*, 1090–1097.
- Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 201–211.
- Sutterer, D. W., Foster, J. J., Adam, K. C. S., Vogel, E. K., & Awh, E. (2019). Item-specific delay activity demonstrates concurrent storage of multiple active neural representations in working memory. *PLOS Biology*, *17*, e3000239.
- Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *Journal of Neuroscience*, *20*, RC63.