Brief Communications

The Head of the Table: Marking the "Front" of An Object Is Tightly Linked with Selection

Yangqing Xu and Steven L. Franconeri

Department of Psychology, Northwestern University, Evanston, Illinois 60208

Objects in the world do not have a surface that can be objectively labeled the "front." We impose this designation on one surface of an object according to several cues, including which surface is associated with the most task-relevant information or the direction of motion of an object. However, when these cues are competing, weak, or absent, we can also flexibly assign one surface as the front. One possibility is that this assignment is guided by the location of the "spotlight" of selection, where the selected region becomes the front. Here we used an electrophysiological correlate to show a direct temporal link between object structure assignments and the spatial locus of selection. We found that when human participants viewed a shape whose front and back surfaces were ambiguous, seeing a given surface as front was associated with selectively attending to that location. In Experiment 1, this pattern occurred during directed rapid (every 1 s) switches in structural percepts. In Experiment 2, this pattern occurred during spontaneous reversals, from 900 ms before to 600 ms after the reported percept. These results suggest that the distribution of selective attention might guide the organization of object structure.

Introduction

Objects in the world do not have objective structural designations, such as axes or orientations. Instead, we assign structural organization to objects as part of visuospatial processing. One fundamental designation within object structure is how we mark one object surface as being its "front." This designation is typically guided by several cues, including which surface is closest to the observer, which surface is most task-relevant or salient (e.g., the face is the front of a head), the source of action (e.g., the infrared emitter is the front of a remote control), or the direction of motion of an object (e.g., the leading edge is the front of a skateboard). However, when these cues are competing, weak, or absent (e.g., if the skateboard stops moving), we can also flexibly assign one surface as the front. For example, Figure 1a depicts a box, a cheese grater, and a table, each with competing surfaces that might be labeled the front. Here we explore a surprisingly simple mechanism that may allow the visual system to flexibly represent this core component of object structure. This abstract structural assignment may be guided by the location of the "spotlight" of selection, where the selected region becomes the front.

Past studies are consistent with the idea that the distribution of selection is associated with organizing object structure. In the ambiguous duck/rabbit illusion, attending to the front of the duck leads to a percept of a duck, and vice-versa for the rabbit (Tsal and Kolbert, 1985). Front assignments can even lead to

changes in relative depth designations, such as in the Necker cube illusion. Recent work using fMRI shows that perceiving a given surface as the front of a Necker cube led to similar activation patterns as a control task where participants selectively attended to that surface of the cube (Slotnick and Yantis, 2005). However, because of the temporal resolution of fMRI, this study could only reveal a coarse temporal link between the two processes.

In the present study we show that seeing a surface as the front has a tight temporal link to the position of the attentional spotlight, using an electrophysiological correlate of spatial selection that allows high temporal resolution measurement. Past studies show that selective processing in one visual field is typically accompanied by a negative deflection at posterior electrodes contralateral to that hemifield [e.g., N2pc, CDA/SPCN (contralateral delay activity/sustained posterior contralateral negativity); for review, see Luck, 2011; Perez and Vogel, 2011]. This technique allows tracking of the distribution of attention without creating dual-task requirements for participants (e.g., probe reporting techniques). In the first experiment, participants were directed to rapidly switch their percepts between the two possible structures of a modified Necker cube. In the second experiment, participants perceived the structural switches spontaneously. We found that shifting of spatial selection was associated with the perceived front surface of an ambiguous figure during both directed and spontaneous switches of percepts.

Materials and Methods

Participants

Eighteen participants (11 female; age range, 18 –35 years) completed Experiment 1, and 13 participants (9 female; age range, 18 –35 years) completed Experiment 2. All participants had normal or corrected-to-normal vision, were paid for participation, and gave written consent.

Stimuli and apparatus

The experiments were controlled by a Dell Precision M65 laptop computer running SR Research Experiment Builder on Windows XP. The
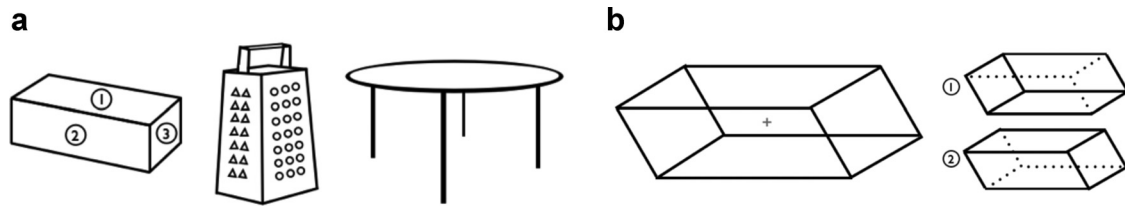
Figure 1. ***a***, For some objects, the location of the front is ambiguous. For the box, one can impose at least three organizations: surface 1 being the front (e.g., a remote control), surface 2 being the front (e.g., a clock radio), or surface 3 being the front (e.g., a security camera). For the cheese grater, the front is determined by the given task. For the table, we can designate an arbitrary position as head of the table. In these cases, the front of an object may be guided by the position of the attentional spotlight. ***b***, In the ambiguous figure task, the display is a modified version of a Necker cube. The cubes on the bottom of the figure depict the unambiguous version of the two possible percepts.

display subtended 32.6° × 24.4° at an approximate viewing distance of 56 cm, on a ViewSonic E70fB CRT monitor with a 75 Hz refresh rate, and 1024 × 768 pixel resolution, 33.6 pixels per degree. Head position was not restrained so as to reduce muscle artifact in the electroencephalographic recording.

The task display was an elongated Necker cube (Fig. 1*b*). The image consisted of two squares, rotated 30° counterclockwise or counterclockwise (counterbalanced and randomized, Fig. 1*b* depicts the counterclockwise version), connected by four horizontal lines, all drawn with light gray lines (31 cd/m^2) against a dark gray (13 cd/m^2) background. The length of the edges of the diamonds was 3.3° and the length of the horizontal edges of the cube was 8.4°.

Procedure
Experiment 1. Before starting the experiment, all participants were given fixation training using a flickering pattern that "jumps" when fixation is broken, which has been shown to improve fixation performance (Guzman-Martinez et al., 2009). Participants pressed a button on the gamepad to initiate a trial. Each trial began with a 500 ms fixation followed by another 800–1200 ms (rectangular distribution) fixation to minimize the impact of previous trials on the EEG signal. A black circle or triangle then surrounded the fixation for 500 ms, and this cue indicated which percept participants should try to perceive first. For example, a circle might indicate that the participant should try to perceive cube (1) and then cube (2) (Fig. 1*b*), while a triangle would indicate the opposite ordering. The words "left" and "right" never appeared in the instructions (only images similar to those in Fig. 1*b*), and response buttons were always vertically arranged. The ambiguous cube was then displayed for 2 s (Fig. 2). An auditory click was played at stimulus onset to prompt participants to perceive the cube in the first structural organization (either (1) or (2)), and then again 1 s later to prompt participants to perceive the alternative organization. At the end of each trial, a response prompt screen appeared. Participants confirmed whether they perceived the cube in the instructed ordering and timing. The two versions of the cube (clockwise or counterclockwise rotated squares, see Stimuli and apparatus, above) and the order of the perceptual organizations were both counterbalanced within participant. Instruction mappings (circle/triangle) were counterbalanced between participants.

Eye movements were monitored by a table-mounted SR Research Eyelink 1000 Remote eyetracker. If participants moved their eyes outside of a 1° radius around the fixation point, from the time window starting from 800–1200 ms preceding the cue (depending on the randomly chosen intertrial jitter value) until their button response, the trial was rejected. Given the small amount of noise present in the eyetracker's position signal (~0.5° root-mean-square error), the effective size of the allowed window was actually smaller than the permitted 1° radius. On rejection, the participant was presented with a screen depicting the allowed fixation region and a dot showing real-time eye position. There was also an indicator of whether the participant had looked left, right, or blinked. The experimenter could then choose to recalibrate the eyetracker at her discretion. The trial was then repeated at a randomly chosen point within the block.

There were a total of 2 blocks of 40 trials. Each block was over when the participant finished 40 eye-movement-free trials. Self-timed

breaks were given after each of these 40 trial blocks. The entire experiment lasted ~150 min, including ERP cap preparation, breaks, and task practice.

Experiment 2. Experiment 2 was similar to Experiment 1, with the following differences. The ambiguous cube was displayed for 8 s, during which the participants were asked to press a corresponding button (vertically arranged on a gamepad) each time their percept changed, while at all times maintaining fixation. Participants were also told to press a third button if their current percept was ambiguous. The entire experiment lasted ~120 min, including ERP cap preparation, breaks, and task practice. On average, participants finished 121 trials within the experiment session.

EEG recording and analysis
Experiment 1. EEG signals were recorded using a BioSemi Active 2 EEG/ERP system. The DC recording was made at 512 Hz with a hardware low-pass filter, and then was decimated in software to 128 Hz. All sites were re-referenced to the postrecording average of the left and right mastoids and high-pass filtered at 0.05 Hz (half-amplitude cutoff). We recorded from 64 silver/silver chloride electrodes mounted in an elastic cap, and the Horizontal and Vertical EOG. The placement of the 64 channels was in accordance with a modification of the international 10/20 system (Jasper, 1958). Two participants were removed from the analysis due to excessive electrode or α noise. For the remaining participants, no trials were rejected. The average HEOG (horizontal EOG) signals for the remaining participants showed a difference of 1.08 μV between percepts (1) and (2), confirming that participants did not systematically move their eyes toward either the perceived front or back of the object (at most a small fraction of a degree; Hillyard and Galambos, 1970; Lins et al., 1993).

Trials with unsuccessful perceptual switches were rejected before the analysis. The EEG data were epoched within a stimulus-locked time window spanning 1 s before the stimulus onset until the button response, and baseline corrected to the 200 ms precue period, which is 700–500 ms before stimulus onset. Electrode PO7/8 were chosen as the electrode-based region of interest based on prior research on N2pc and CDA (for review, see Luck, 2011; Perez and Vogel, 2011).

Experiment 2. Experiment 2 was similar to Experiment 1, with the exception that we recorded from the following sites according to the 64-channel modification of the international 10/20 system: F3/4, C3/4, PO3/4, P5/6, P7/8, PO7/8, O1/2, POz, Oz, Horizontal and Vertical EOG. Two participants were removed from the analysis due to noisy electrodes. The average HEOG signals for the remaining participants showed a difference of 0.35 uv between percepts (1) and (2), confirming that participants did not systematically move their eyes toward either the perceived front or back of the object (at most a small fraction of a degree; Hillyard and Galambos, 1970; Lins et al., 1993). The EEG data were epoched within a response-locked time window spanning 2 s before and 2 s after the report of a perceptual change, and baseline corrected to the 200 ms prestimulus period. Epoch boundaries of the 4 s window were trimmed to exclude time ranges corresponding to the alternative percept or ambiguous percepts.
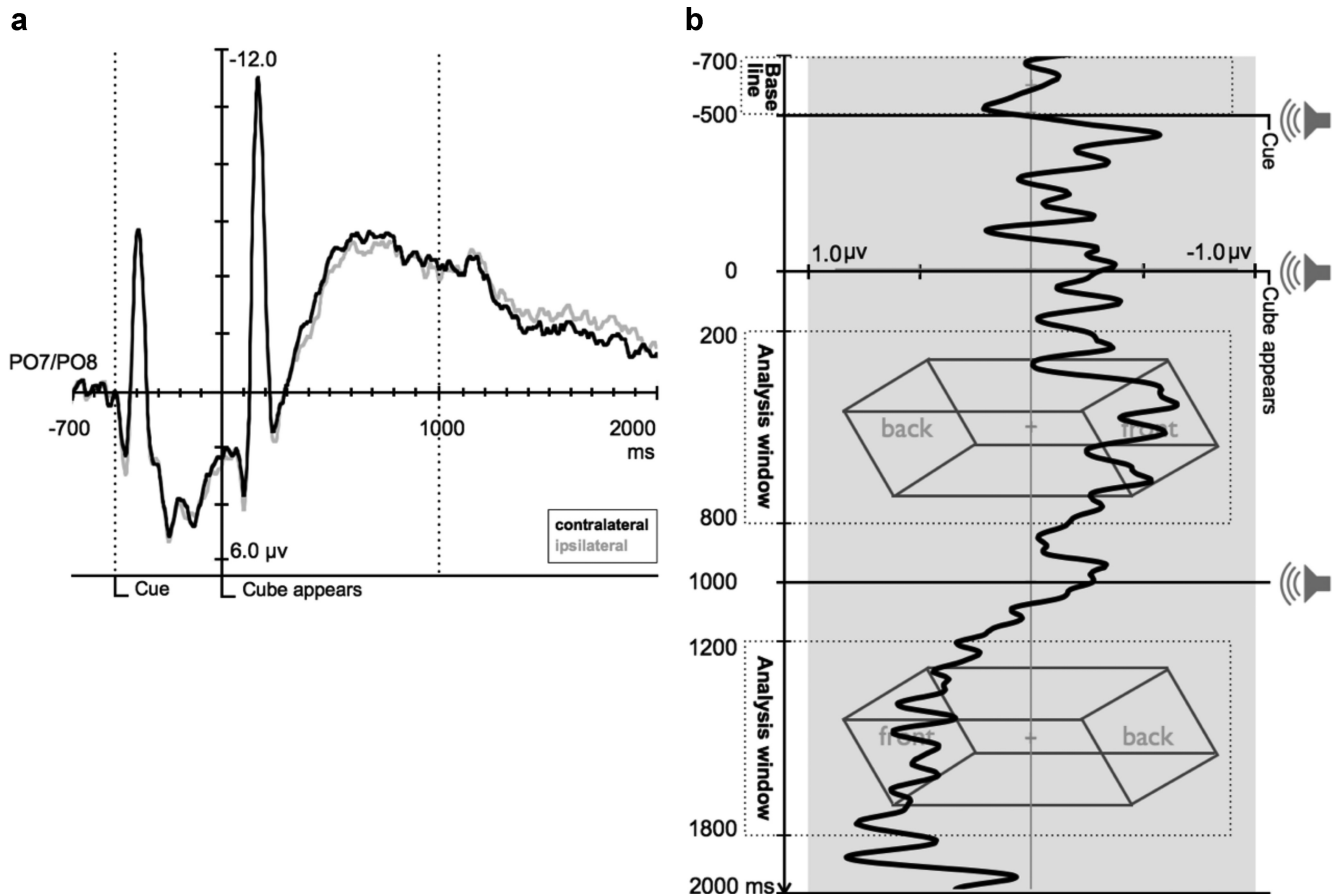
a

b

Figure 2. *a*, The grand average ERP waveforms at PO7/PO8, ipsilateral and contralateral relative to the first perceived front side, time locked to the onset of the ambiguous cube. The activity visible in the prestimulus period is due to the response evoked by the cue display. *b*, The difference waveforms in the ambiguous figure task, indicating relative selection of the front or back side of the cube. After a 200 ms baseline period, a shape cue informs participants of the temporal order in which they should perceive the cube switch in structure across the two time windows (e.g., in the sample below, left-back/right-front, then left-front/right-back). After 500 ms preparation time, the cube appears for 2 s. The difference waveform shows that during the first time window participants shift selection toward the perceived front of the cube, and in the second time window they shift selection to the new perceived front of the cube. The magnitude of the difference wave does not reflect the actual position of spatial selection on the screen. For visual clarity, all waveforms were low-pass-filtered by convolving them with a Gaussian impulse–response function (SD = 16 ms; 50% amplitude cutoff at ∼20 Hz).

Results

Experiment 1: ERP correlates of selection and directed switches of percepts

The average success rate for perceiving the cube in the instructed ordering and timing was 78.6% across all participants. Figure 2*a* depicts the raw contralateral potentials (the average of PO7 for percept (2) ("right-front-left-back") and PO8 for percept (1) ("left-front-right-back") and ipsilateral potentials (the average of PO7 for percept (1) (left-front-right-back) and PO8 for percept (2) (right-front-left-back) averaged across all participants. Figure 2*b* depicts the difference waveform, as the subtraction of the ipsilateral waveform from the contralateral waveform. Two measurement time windows of 600 ms were extracted from each epoch (200–800 ms, and 1200–1800 ms after cube onset). This time window was chosen a priori, based on (1) the earlier component of interest (N2pc) begins after 200 ms (for review, see Luck, 2011), (2) the 800 ms offset was chosen for symmetry, (3) participants likely begin anticipatory shifts of selection toward the other side.

As seen in the results depicted in Figure 2*b*, the difference wave was further to the right (more negative for electrodes contralateral to the perceived front side of the figure) in time window 1 relative to time window 2 when they perceived the other side as the front, reflecting preferential selection of the currently perceived front side of the cube. Potentials were more negative for electrodes contralateral

to the perceived front side of the figure in the first time window [mean (M) = -0.38 μV], relative to those same electrodes in the second time window that were now contralateral to the figure's back (M = 0.49 μV), $t_{(15)} = 2.57$, $p = 0.021$. Comparing the absolute values of these windows did not reveal that this preference was stronger in either time window, $t_{(15)} = 0.268$, $p = 0.792$, n.s. Figure 2*b* also shows a small "bump" immediately after the cue (-500 ms) that could indicate an anticipatory shift of selection toward the future position of the object front. However, this trend was not significant—the difference wave for the time window -300 ms through stimulus onset (the typical N2pc/CDA time range) did not differ significantly from zero (M = -0.13 μV), $t_{(15)} = 0.806$, $p = 0.433$, n.s. Figure 3 depicts the topographic voltage map of the difference wave.

In summary, these results indicate that shifting of spatial selection was associated with the perceived front surface of an ambiguous figure when participants were told to perceive the figure in different structures.

Experiment 2: ERP correlates of selection and spontaneous switches of percepts

The average onset of the first percept in one trial was ∼1.5 s. Subsequent reversal rate was 2.9 s/switch on average. On average, participants pressed the button corresponding to percept (1) 0.95
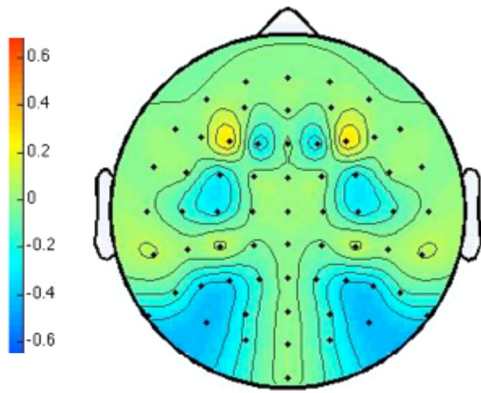
Figure 3. The topographic voltage map of the difference wave. We computed the potentials contralateral to the perceived front minus those ipsilateral to the perceived front in the first time window (200–800 ms) and the potentials contralateral to the perceived front minus those ipsilateral to the perceived front in the second time window (1200–1800 ms). The topography depicts the average of these two contralateral minus ipsilateral differences, collapsed across left and right, and mirrored across the midline.
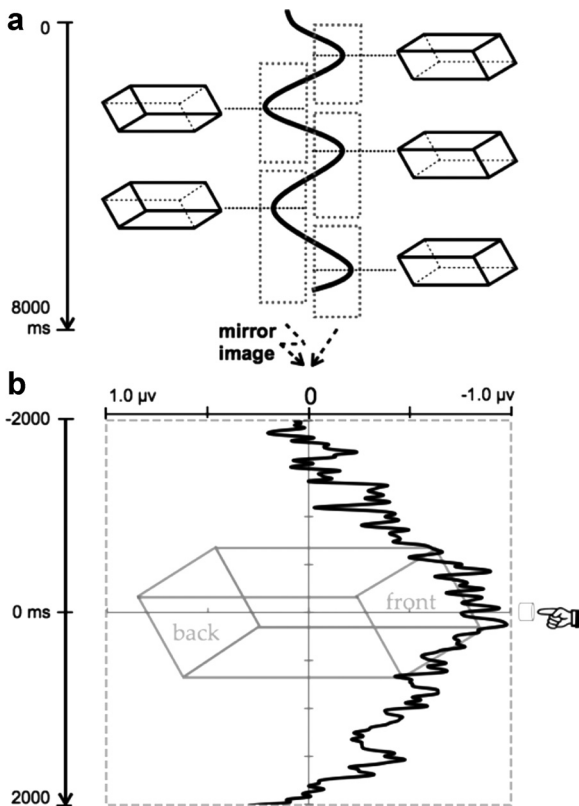

Figure 4. *a*, A schematic version of the analysis technique. Within an 8 s trial there could be several reports of a perceptual switch in the structure of the cube. We took response-locked ERPs at each report of a switch (see Materials and Methods for details), and collapsed the two types of percept reports into a difference wave showing activity contralateral to the new perceived front of the cube. *b*, The grand average of this difference wave across subjects. The magnitude of this difference wave does not reflect the actual position of spatial selection on the screen. The results show more PO7/PO8 negativity contralateral to the front face of the cube 900 ms before and 600 ms after the switch report, suggesting a shift of attention toward the new front side.

times and the button corresponding to percept (2) 1.17 times in each trial. Figure 4*b* depicts the difference waveform as the subtraction of the ipsilateral waveform from the contralateral waveform. That is, the average of PO7 for percept (2) (right-front-left-back) and PO8 for percept (1) (left-front-right-back)

minus the average of PO7 for percept (1) (left-front-right-back) and PO8 for percept (2) (right-front-left-back). The amplitude of the difference wave (M = −0.42 μV) was significantly less than zero in the time period spanning from 2 s before and 2 s after the response, $t_{(10)} = 3.04$, $p = 0.01$. The data were then analyzed in a repeated-measures ANOVA with 40 measurement time windows of 100 ms (2 s before to 2 s after the response). There was a significant main effect of time, $F_{(39,290)} = 1.898$, $p = 0.001$, $\eta^2 = 1.6$, indicating differences in difference wave magnitude across time. To determine the time point at which the contralateral potentials were significantly more negative than the ipsilateral potentials (i.e., the time at which the difference wave becomes significantly different from zero), we performed a t test of the two electrode sites at each time point. We adjusted for multiple comparisons by finding the time point that meets the following two criteria: (1) the p-value was <0.05, and (2) the p-values for the subsequent 100 ms were all <0.05 (Luck, 2005). Using this method, we found that the difference wave was significantly different from zero in the time window starting 891 ms before the response until 609 ms after the response, all p-values <0.045, all t-values >2.29. These results suggested that, on average, at ~900 ms before reporting a perceptual switch, participants had selected the side of the figure that they were about to see as the front of the figure.

Discussion

The present study shows a tight temporal link between selective attention and the perceived structural organization of an ambiguous figure, during both directed and spontaneous changes of percepts. In the first experiment, participants were directed to switch their percepts at a fixed tempo. During the first time window, participants shifted selection toward the perceived front of the cube, and in the second time window they shifted to the new perceived front of the cube. In the second experiment, participants switched their percepts spontaneously. The results show that on average participants shifted to the new front of the cube for the time period 900 ms before and 600 ms after the switch report.

The contralateral negative signal observed in our experiments is likely a mixture of N2pc, a transient negativity 175–300 ms poststimulus in posterior areas contralateral to the attended visual hemifield, and CDA/SPCN, a similar component which appears 275–900 ms poststimulus. The N2pc component is argued to reflect selection of stimuli in one visual field (Luck, 2011), while the CDA/SPCN is argued to reflect continuous processing or memory encoding of stimuli in one visual field (Perez and Vogel, 2011). For the purpose of the present study, it is sufficient to consider both components as reflecting selective processing of areas within one visual field relative to the other visual field.

While our results demonstrate a close temporal link between spatial selection and perceived object structure, the present data cannot by themselves demonstrate that selection causes changes in structure. One may argue instead that observers first assign structure to an object and then shift selection to the currently perceived front. However, we speculate that selection plays a causal role. First, in Experiment 2 there was evidence of a shift 900 ms before the change in percept, though this evidence carries the caveats that (1) it is difficult to know how much of this effect is due to variability in report timing, and (2) the 800 ms is before the report of the percept, but the percept must have occurred shortly before the report. Second, previous behavioral studies have found that spatial selection can cause changes to interpretations of ambiguous images (Tsal and Kolbert, 1985; Peterson and Gib-

son, 1991). Manipulating an observer's initial gaze can also bias percepts of the Necker cube (Ellis and Stark, 1978) and the ambiguous old lady/young lady illusion (Georgiades and Harris, 1997). Similar effects also occur within the interpretation of ambiguous scenes. In one study, participants were presented with images depicting events that can be described in either an active way or a passive way (e.g., "The dog is chasing the man." or "The man is running away from the dog.") Directing participants' attention to different locations of the scene altered the sentence structure and the word choice when the scene was later described (Gleitman et al., 2007).

The locus of spatial selection might serve to guide object structure because the features that drive a surface to be designated the front—the closest surface, most task-relevant or salient surface, the source of action, or the direction of motion—are all features that also reliably cause a surface to be selected. Through experience, this relationship could serve to train a correlation between selection and the front. This correlation could then serve as a tool that allows an observer to control their interpretation of an object's structure. Attending to one surface (He and Nakayama, 1995) could increase the cortical response associated with that area, biasing competition between potential structural representations (Desimone and Duncan, 1995; Luck et al., 1997), and guiding or maintaining the corresponding structure (Slotnick and Yantis, 2005). This role would promote selection from a filter to a more sophisticated agent in creating complex visual representations (Ullman, 1984; Peterson and Gibson, 1991; Cavanagh, 2004; Franconeri et al., 2012).

Using spatial selection as a marker for abstracted object structure could supplement other systems for object recognition. Hummel and Biederman (1992) argue that object recognition could be accomplished by a multilayer network containing progressively more sophisticated layers of processing. However, this network does not code the direction of components (e.g., left-pointing horizontal vs right-pointing horizontal). The mechanism we describe here could supplement such models by allowing the specification of a direction for a specific component. Selection might also play a role in guiding dynamic processing of object structure, such as mental rotation. Rotating an object in one's mind's eye requires the observer to store a structural representation of the object in working memory. It is still unclear how exactly such structural representation is created and controlled in mental rotation (for review, see Zacks, 2008). It is possible that the spotlight of spatial attention can mark certain surface as special relative to others, which enables the visual system to "push" the object flexibly around an axis during mental imagery.

References

Cavanagh P (2004) Attention routines and the architecture of selection. In: Cognitive neuroscience of attention (Michael Posner, ed), pp 13–28. New York: Guilford.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

Ellis SR, Stark L (1978) Eye movements during the viewing of Necker cubes. Perception 7:575–581.

Franconeri SL, Scimeca JM, Roth JC, Helseth SA, Kahn LE (2012) Flexible visual processing of spatial relationships. Cognition 122:210–227.

Georgiades MS, Harris JP (1997) Biasing effects in ambiguous figures: removal or fixation of critical features can affect perception. Vis Cogn 4:383–408.

Gleitman LR, January D, Nappa R, Trueswell JC (2007) On the give and take between event apprehension and utterance formulation. J Mem Lang 57:544–569.

Guzman-Martinez E, Leung P, Franconeri S, Grabowecky M, Suzuki S (2009) Rapid eye-fixation training without eye tracking. Psychon Bull Rev 16:491–496.

He ZJ, Nakayama K (1995) Visual attention to surfaces in 3-D space. Proc Natl Acad Sci U S A 92:11155–11159.

Hillyard SA, Galambos R (1970) Eye movement artifact in the CNV. Electroencephalogr Clin Neurophysiol 28:173–182.

Hummel JE, Biederman I (1992) Dynamic binding in a neural network for shape recognition. Psychol Rev 99:480–517.

Jasper HH (1958) The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol 10:371–375.

Lins OG, Picton TW, Berg P, Scherg M (1993) Ocular artifacts in recording EEGs and event-related potentials. II: Source dipoles and source components. Brain Topogr 6:65–78.

Luck SJ (2005) An introduction to the event-related potential technique. Cambridge, MA: MIT.

Luck SJ (2011) Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP components. In: The Oxford handbook of event-related potential components (Luck SJ, Kappenman ES, eds), pp 329–360. New York: Oxford UP.

Luck SJ, Girelli M, McDermott MT, Ford MA (1997) Bridging the gap between monkey neurophysiology and human perception: an ambiguity resolution theory of visual selective attention. Cogn Psychol 33:64–87.

Perez VB, Vogel EK (2011) What ERPs can tell us about visual working memory. In: The Oxford handbook of event-related potential components (Luck SJ, Kappenman ES, eds), pp 361–372. New York: Oxford UP.

Peterson MA, Gibson BS (1991) Directing spatial attention within an object: altering the functional equivalence of shape descriptions. J Exp Psychol Hum Percept Perform 17:170–182.

Slotnick SD, Yantis S (2005) Common neural substrates for the control and effects of visual attention and perceptual bistability. Cogn Brain Res 24:97–108.

Tsal Y, Kolbert L (1985) Disambiguating ambiguous figures by selective attention. Q J Exp Psychol 37(A):25–37.

Ullman S (1984) Visual routines. Cognition 18:97–159.

Zacks JM (2008) Neuroimaging studies of mental rotation: a meta-analysis and review. J Cogn Neurosci 20:1–19.