Current Biology *17*, 26–31, January 9, 2007 ©2007 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2006.10.050

Report

Interference with Bottom-Up Feature Detection by Higher-Level Object Recognition

Li Zhaoping[1,*] **and Nathalie Guyader**[1]
[1] Department of Psychology
University College London
London WC1E 6BT
United Kingdom

Summary

Drawing portraits upside down is a trick that allows novice artists to reproduce lower-level image features, e.g., contours, while reducing interference from higher-level face cognition. Limiting the available processing time to suffice for lower- but not higher-level operations is a more general way of reducing interference. We elucidate this interference in a novel visual-search task to find a target among distractors. The target had a unique lower-level orientation *feature* but was identical to distractors in its higher-level *object shape.* Through bottom-up processes, the unique feature attracted gaze to the target [1–3]. Subsequently, recognizing the attended object as identically shaped as the distractors, viewpoint invariant object recognition [4, 5] interfered. Consequently, gaze often abandoned the target to search elsewhere. If the search stimulus was extinguished at time T after the gaze arrived at the target, reports of target location were more accurate for shorter (T < 500 ms) presentations. This object-to-feature interference, though perhaps unexpected, could underlie common phenomena such as the visual-search asymmetry that finding a familiar letter N among its mirror images is more difficult than the converse [6]. Our results should enable additional examination of known phenomena and interactions between different levels of visual processes.

Results and Discussion

Among the 45° left-tilted bars in Figure 1, a uniquely right-tilted bar, 45° or 20° from the vertical in conditions A_{simple} or B_{simple}, pops out. However, superposing a horizontal or vertical bar on each original bar makes the uniquely tilted bar much harder to find in condition A than condition B of Figure 1. The target object in condition A but not B is a rotated and sometimes also a mirror-reversed version of all distractor objects, easily confused with the distractors because object recognition is typically rotationally or viewpoint invariant. We suggest that the higher-level perception of the object comprising the two intersecting bars interferes with the task of locating it based on its unique lower-level orientation feature component.

Primitive features, like the orientations of small bars, of visual inputs are first extracted by the primary visual cortex (V1) [7]. Then these features are combined into objects, e.g., composed of two intersecting bars [8, 9], by higher cortical areas, including the inferotemporal (IT) cortex, whose neurons are selective to object shapes [10–15]. V1 is not only a way station; its activities also highlight salient items because of its sensitivity to unique low-level features such as orientation [16–18]. In addition to driving the higher visual areas such as V4, which combines bottom-up and top-down factors [19–21], V1's saliency signal also evokes cognitive decisions by driving superior colliculus, which controls saccades [3]. Behaviorally and preattentively, unique image features such as orientation and color can pop out [1], and an object's basic features such as "vertical" and "red," but not its overall shape, can be obtained [22]. Meanwhile, an important characteristic of the progression from feature to object processing is making object recognition viewpoint independent [23] and thus achieving object invariance. Some IT neurons are indeed insensitive to viewpoint [10–12]. IT activities also correlate with the planning of saccades [24]. There is thus a hierarchy of levels of cognition and their consequent decisions and actions. Behaviorally, attentive exposure to an object's image can prime its subsequent recognition regardless of viewpoint but can prime its recognition only in the same view if the exposure was unattended [4].

The observations above suggest the following relevant processing stages: (1) an early preattentive stage that processes image features, e.g., orientations of object components, and makes unique features salient [1]; and (2) a later, attentive [4] stage that creates a viewpoint-invariant object representation [1, 5], e.g., a shape from two intersecting bars. For locating a target possessing a uniquely oriented bar in the display, the early stage suffices because the salient unique orientation can attract gaze. The later, attentive object-processing stage is commonly expected to facilitate processing of the components of the objects through top-down feedback [25]. However, when differently oriented but otherwise identical distractors are present, as in condition A but not B of Figure 1, viewpoint-invariant object recognition could make search harder. If so, briefer stimulus viewings (within a time window), preventing invariant object recognition, should improve target localization in condition A but not B. We show exactly this below.

In experiment I, subjects searched among 660 objects, in a display extending 46° × 34° of visual angle, for the object with the uniquely tilted oblique bar. The search stimulus was of conditions A_{simple}, B_{simple}, A, or B (Figure 1) or control conditions. The nonoblique, task-irrelevant bar in the target of condition A or B was randomly either horizontal or vertical (the task-relevant bar in condition B was always 20° from this irrelevant bar). The subjects were a priori informed about the uniquely oriented target bar and that this unique orientation could be randomly tilted to the left or right in each trial. They were asked to press a left or right button quickly to indicate whether the target was in the left or right half of the display. Their eye positions were tracked.

---

*Correspondence: z.li@ucl.ac.uk

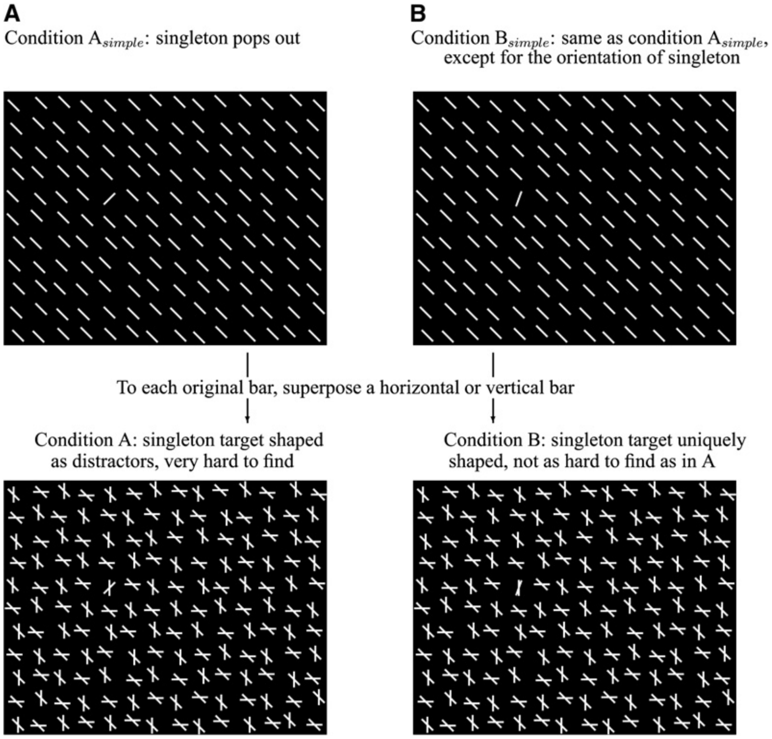

Figure 1. Small Portions of Visual-Search Displays

The target possesses the uniquely left- or right-tilted (as in these examples) bar in the entire display. In conditions A_{simple} and B_{simple} (top), all bars were 45° from vertical, except the target bar in B_{simple}, which was 20° from vertical (in this example) or horizontal. Conditions A and B (bottom), derived from A_{simple} and B_{simple}, differed only in the angle, 45° and 20°, respectively, between the two bars in the target. Task-irrelevant, horizontal and vertical, bars made the orientation singleton much harder to find in condition A than in condition B.

Figure 2 shows that reaction times (RTs) for the subject's first gaze arrival to the target, RT_{eye}s, were comparable in conditions A and B. This is unsurprising because the target in both conditions had the uniquely oriented bar. This bar is salient preattentively [1, 2], attracting both attention and gaze, the latter because of the mandatory link between the directions of attention and gaze in free viewing [26]. These RT_{eye}s were longer than those in conditions A_{simple} and B_{simple} mainly because the nonuniform orientations of the task-irrelevant (horizontal and vertical) bars reduced the target's saliency [27]. However, the RTs for reporting the target location by button press, RT_{hand}s, were typically more than 1−2 s longer in condition A than B, even though A and B had comparable button-response accuracies. In condition A, after gaze first reached the target, it often dawdled

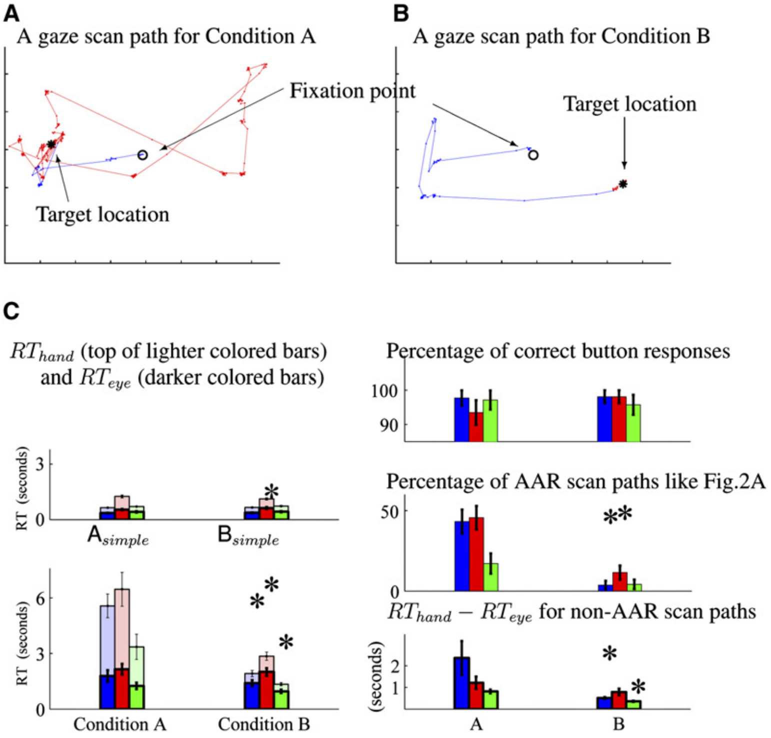

Figure 2. Hand and Gaze Responses in Experiment I

(A and B) Examples of gaze scan paths. The one in (A) is for an arrive-abandon-return (AAR) trial. Asterisks and open circles mark the locations for targets and fixation points, respectively; the grid frames the spatial extent of the stimuli. Blue and red scan paths, respectively, are for those before and after the first gaze's arrival to the target and before the button press.

(C) Data for three subjects, denoted by red, green, and blue hues, respectively. Asterisks indicate significant differences between conditions for the subject. The left graphs show RT_{hand} and RT_{eye} (top of lighter- and darker-colored bars, respectively) for button responses and first gaze at target, respectively. The right graphs show task performances, percentage of arrive-abandon return (AAR) scan paths (e.g., [A]), and eye-to-hand latencies in non-AAR trials for conditions A and B only. All error bars show SEM.

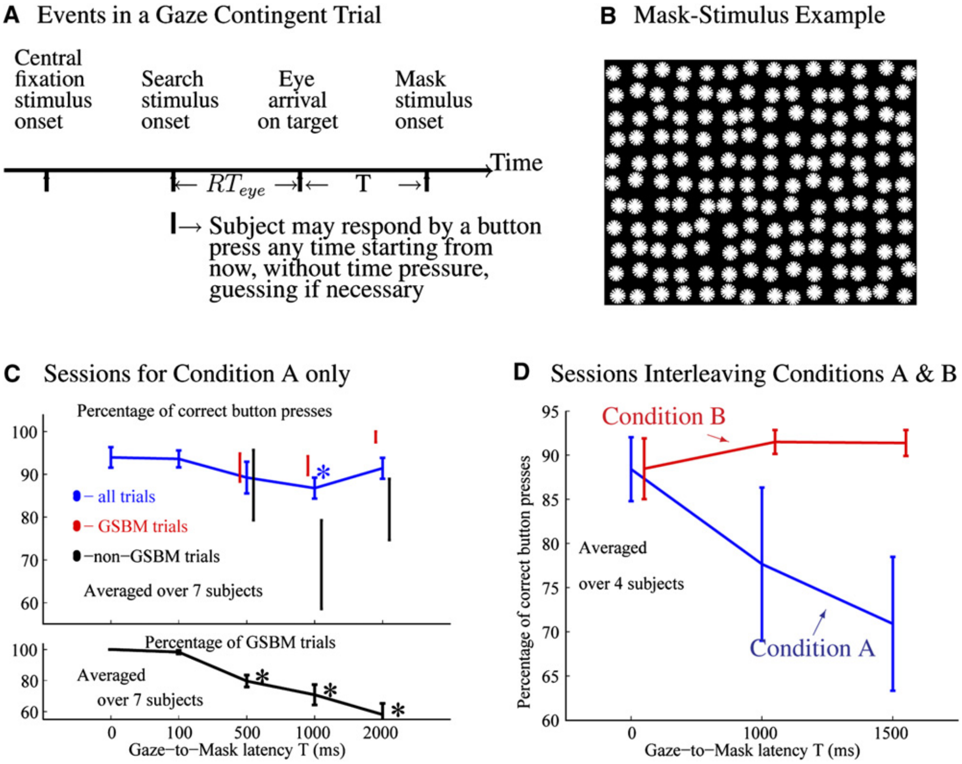

Figure 3. Experiment II: The Longer One Looks, the Worse One "Sees"

(A) Sequence of events in a gaze-contingent trial.

(B) A small portion of an example of a mask stimulus.

(C) With longer gaze-to-mask latency T, target localization in condition A (in blocked sessions) worsened, and the gazes are more likely to have abandoned the target before mask onset. Asterisks denote data points significantly smaller in value than that for T = 0. GSBM trials are those in which *gaze stayed (at target) before mask* onset.

(D) In sessions interleaving conditions A and B (for another subject group), performances in conditions A and B are comparable for T = 0. Combining both T > 0 values, performance in B is significantly better than that in A ($p = 0.01$). Error bars show SEM.

around the target before the button press or even abandoned the target to search elsewhere before returning to it prior to the button press (Figure 2A). Such arrive-abandon-return (AAR) scan paths were much rarer in condition B. Even for the non-AAR trials, the eye-to-hand latency $RT_{hand} - RT_{eye}$ was much longer in condition A than in B. These observations are consistent with the hypothesis that decision processes vetoed the first guess by the feature detectors in condition A because the attended object was recognized as having the same shape as the distractors, i.e., invariant object recognition could be interfering.

An alternative explanation consistent with the data could be that somehow targets in condition A but not condition B become less visible under foveal viewing than peripheral viewing. To test between the hypotheses of interference by invariant object recognition and of foveal visibility reduction, we examined conditions A and B in experiment II, in which the search stimulus was masked after a seemingly random time interval since its onset (Figure 3A). The subjects button-pressed for the target location as before, but could respond without time pressure, before or after the mask onset, and guess if they had to. The mask (Figure 3B) covered each original object, whether target or distractor, with a star-shaped object and made the original object imperceptible. A random half of the trials in each session were gaze-contingent trials, in which mask onset occurred and reduced visibility of the original stimulus to zero at one of several predetermined time intervals T after gaze first arrived at the target. The other trials had random mask-onset times; some were gaze-opposite trials, in which the gaze position at mask onset was on one (e.g., left) side the display center and the target was on the opposite (e.g. right) side of the display center, and were designed to prevent subjects' awareness of any link between mask onset and eye position (see Experimental Procedures).

Figure 3C shows that for condition A, target localization worsened with longer gaze-to-mask viewing time $T \leq 1$–2 s. This is not because the button presses tended to agree with the eye positions at mask onset; among the gaze-opposite trials, only 56% of the button presses agreed with the eye positions at mask onset. Furthermore, the performance for T = 0, when target visibility became zero immediately upon foveal viewing, is comparable to that without the mask in experiment I, when the stimulus was viewed as long as was deemed necessary by the subjects. This suggests that the extra viewing time T > 0, or a longer duration of target visibility (even if reduced), is unnecessary and can be detrimental for target localization for some T. Apparently, the subjects had a good first guess of the target location based on image features (orientations of the bars) alone before

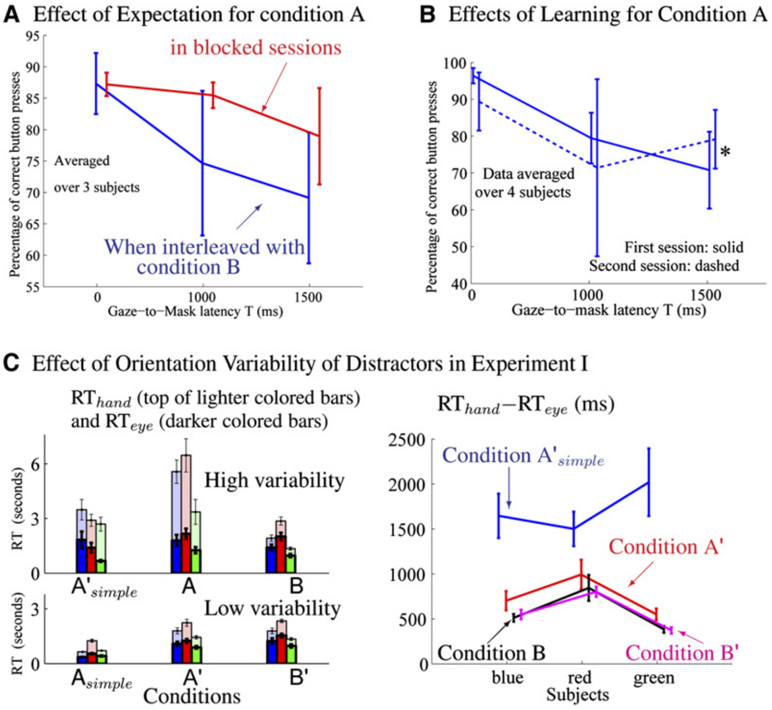

Figure 4. Factors Affecting Object-to-Feature Interference

(A) In experiment II, performance across $T > 0$ for condition A was somewhat better ($p = 0.08$) in blocked sessions (one session each subject) than in sessions interleaved with condition B (two sessions each subject), in which subjects had higher expectations for uniquely shaped target.

(B) Reduction of interference with experience—for longer gaze-to-mask time T in experiment II, performance for condition A improved in the second experimental session (significantly at T value with an asterisk next to the data points). The data in Figure 3D were replotted here according to the two separate sessions.

(C) Stronger or weaker object-to-feature interference, manifested in $RT_{hand} - RT_{eye}$ in experiment I, by, respectively, higher or lower orientation variabilities of the distractors for reducing or enhancing bottom-up pop-out strength manifested in RT_{eye} (same three subjects as those in Figure 2, denoted by blue, red, and green colors). The $RT_{hand} - RT_{eye}$ in condition A′, although reduced from that of A, is significantly longer ($p = 0.002$) than that of B′. In the right graph, data points of different conditions are plotted in different colors. Stimulus examples of conditions A′, B′, and A'_{simple} are shown in the Supplemental Data. Error bars show SEM.

they got confused by invariant object recognition, which likely caused them to abandon target (the non-GSBM trials in Figure 3C) and give incorrect responses. Eventually, their confusion subsided. Some subjects reported that sometimes they thought they found the target, only for it to disappear when they took a second look. In experimental sessions interleaving conditions A and B (for another group of subjects), extra viewing time $T > 0$ improved performance in condition B marginally but worsened performance in A (Figure 3D). Meanwhile, the fact that the performances for the two conditions at $T = 0$ are comparable is consistent with the comparable RT_{eye}s in these conditions in experiment I (Figure 2).

Our finding is the first we know of providing quantitative psychophysical data to suggest that deeper cognitive processing can be detrimental to some visual cognitive tasks—a likely explanation for the portrait-drawing trick. In particular, invariant object recognition interfered with lower-level feature processes' abilities to detect unique salient features. Here, the later-stage processes for object recognition are at best unnecessary for our task. Our findings suggest that they actually overwrite or interfere with the decisions of the necessary and earlier feature processes, even though, in principle, they do not have to do so. The uniqueness of the orientation of the target's component bar is sufficient to make the target location salient. Previous physiological and computational studies [16–18, 2] have indicated that V1 can detect and highlight such a salient feature and direct gaze to it via the superior colliculus [3].

Although some forms of object recognition can occur quickly [28, 29] and without attention or awareness [5, 30], psychophysical data have indicated that viewpoint-invariant object representation needs attention [4, 22]. Accordingly, our findings suggest that the later, interfering stage does not only construct object from features but also allows top-down attention to build invariant object representations. This is consistent with the mandatory link between the directions of gaze and attention in free viewing [26]. Thus, our finding can also be seen as the interference of top-down attentional processes with bottom-up processes, and this interference introduces nontrivial complexity to the temporal and performance differences between higher- and lower-level processes [31–33]. Our finding also contrasts with backward visual masking [34] in which inattention enables a mask to impair object recognition. Figure 3C suggests that building the invariant object representation requires at least 100 ms of attentive viewing for objects in our stimuli.

Our analysis suggests the following factors as being conducive to interference: (1) tasks being feature based, not requiring object recognition; (2) object recognition or top-down knowledge, or both, introducing additional signals, which has sufficient weight to counteract the low level feature's contribution to task-relevant decisions. Comparing condition A in blocked versus interleaved (with condition B) sessions (Figure 4A) suggests that an increased expectation for a unique target shape (in the interleaved session) increases interference. This is unsurprising because the expectation should increase the weight of factor (2) above. Analogously, we can reduce interference by increasing the weight of the bottom-up factor and thus decreasing the relative weight of the factor (2). For instance, when the task-irrelevant bars in conditions A and B are all horizontal or all vertical so that distractors are uniformly oriented, the target becomes more salient. We call these modified conditions A′ and B′, respectively. This reduces RT_{eye} significantly. Consequently, the feature-level influences could more strongly push the task-decision process so that the decision threshold could be reached before object-to-feature interference becomes more significant. Hence, in experiment I interleaving conditions A′, B′, A, and B, $RT_{hand} - RT_{eye}$ for A′ is much shorter than for A, although $RT_{hand} - RT_{eye}$ for A′ is still significantly longer than the two comparable $RT_{hand} - RT_{eye}$ values for B and B′ (Figure 4C). Conversely, when the orientation variability of distractors is increased in condition A_{simple}, such that randomly 1/3 of the distractor bars become oriented horizontally and another 1/3 become oriented vertically, we call the resulting stimulus condition A'_{simple}. In this condition A'_{simple}, object-to-feature interference arises by a $RT_{hand} - RT_{eye}$ longer than that in condition B (Figure 4C). This suggests that even a simple bottom-up orientation feature can, given sufficient processing time, be treated as a viewpoint-invariant object bar and make the target object bar a rotated version of all distractor objects.

Our data also suggest that subjects can quickly learn to remove the interference in condition A within two data sessions involving no more than 260 trials per subjects in experiment II (Figure 4B). Subjects reported discovering helpful strategies of trusting their instincts, defocusing the image, or letting the target pop out while fixating on the center of display away from the peripheral target. Peripheral visual field is more heavily sampled by the magno cellular pathway, which, compared to the parvo cellular pathway, is faster and processes coarser resolution inputs [35, 36]. Hence, the magno pathway likely plays a greater role in detecting unique features and driving gaze in a bottom-up manner. This is consistent with the idea that slower attentive process is associated with finer spatial resolution than the faster bottom-up processes. Defocusing and peripheral viewing probably reduce the object-to-feature interference by selectively emphasizing the magno pathway to speed up the bottom-up process while removing the finer input details to attenuate the attentive object-formation processes. Although removing finer resolution could make two intersecting bars resemble a single bar of the averaged orientation, the observed object-to-feature interference in condition A'_{simple} (which has only disconnected bar stimuli) suggests that viewing the objects as single bars could not remove the interference if attentive object formation proceeded. Hence, we predict that lesions (clinical or by transcranial magnetic stimulation) of the cortical areas responsible for attentive-object processes (perhaps the parietal cortex, which has been implicated in building objects from features [5]) could improve performance in our task. Our findings only reveal a fraction of the rich interactions between lower- and higher-level cognitive processes. The results of such interactions are unexpected if we assume that deepening of processes should always lead to improved perception.

Different degrees of object-to-feature interference may underlie common observations of visual-search asymmetry between familiar and unfamiliar targets. For example, a search for a familiar letter N among its mirror reversals is performed more slowly than a search for a mirror reversal among normal N's [6, 37, 38]. Both searches require the same low-level processes for detecting orientation contrast between left- and right-tilted bars and do not require letter recognition. However, familiarity of the letters should affect the object rather than feature-level processing. Hence, the object-to-feature interference, manifested in our task and likely behind the portrait-drawing trick, can enable additional examination of many known phenomena.

Experimental Procedures

Stimuli

Each stimulus display, viewed at a distance of 40 cm, had 660 object items, each at a position randomly displaced, up to ±0.24° visual angle, horizontally and vertically from its corresponding position in a regular grid of 22 rows × 30 columns, spanning correspondingly 34° × 46° in visual angle. Each stimulus bar was 0.12° × 1.1° in visual angle and 48 cd (candela)/m^2 in brightness. The background was black. The target's grid location was randomly one of those closest to the circle of about 15° eccentricity, and beyond 12° of horizontal eccentricity, from the display center. The fixation stimulus was a bright disk of 0.3° diameter at the display center.

Procedures

Gazes were tracked by the 50 Hz infrared video eye tracker from Cambridge Research System (www.crsltd.com). Tracking calibration was performed before each data session to a precision typically within 0.5° of visual angle. After being shown two examples of each stimulus condition, untrained subjects were instructed to fixate centrally until the stimulus onset and to freely move their eyes afterwards for target searching. The sequence of events in a trial was as follows: (1) With the fixation stimulus, the subject pressed a button to start a trial and eye tracking. (2) After 0.6 s, upon the subject's continuous fixation for 40 ms within 3° of the fixation point, a blank screen replaced the fixation stimulus for 200 ms and was followed by the onset (designated as time zero) of search stimulus. (3) In experiment I, the search stimulus remained till after the subject's button press. In experiment II, a mask replaced the search stimulus at a time determined as follows: In a gaze-contingent trial, the mask onset occurred at time T after the first gaze arrival at the target. The criterion for the arrival was when the gaze was within 2.3° in visual angle from the target's center position. T was randomly chosen from the set T = (0, 100, 500, 1000, and 2000) ms for data sessions contributing to Figure 3C, and for a different group of subjects, from the set T = (0 and 1500) ms or T = (0, 1000, and 1500) ms for sessions contributing to Figure 3D and Figures 4A and 4B. For each non-gaze-contingent trial, a time τ was chosen randomly and uniformly from the time window 200–1700 ms. The mask onset occurred upon the first gaze arrival at the opposite (laterally from the center) side of the target since 200 ms after stimulus onset or at time τ, whichever was sooner. The mask, once displayed, remained until after the subject's button press. Each session of experiment I had 200 trials, randomly interleaving conditions A_{simple}, B_{simple}, A, B, A′, B′, and A'_{simple} and other control conditions. In experiment II, each blocked session for condition A had 130 or 60 trials, and each interleaving session (of conditions A and B) had 100 trials. After each session of experiment II, we verified that subjects did not notice any links between the mask onsets and the gaze positions. Different subjects participated in experiments I and II.

Data Analysis

A trial is defined as a bad trial and removed from further analysis if gaze was untracked in more than 10% of the video frames of the eye tracker within the time window (0, RT_{hand}) or if $RT_{hand} < 100$ ms. Data from a subject or session when bad trials comprised more than 10% of all trials are removed from further analysis. Sufficiently large gaze-tracking error can lead to failures in detecting gaze arrivals at the target. A trial is called a nonarrival trial if the gaze never arrived at the target by our arrival criteria with the tracker measurements. We thus remove from further analysis subjects and data sessions having more than 11% of nonarrival trials among all trials in experiment I or among the gaze-contingent trials in experiment II. Results in figures were based on the gaze-arrival trials only. The RTs plotted were based on trials with correct button responses. The error bars plotted represent the standard error of the mean (SEM). Statistical tests for differences between different conditions in Figure 2 were by two-tailed t test, whereas those in Figures 3 and 4 were by one-tail matched sample t test.

Supplemental Data

Supplemental Data include additional Experimental Procedures and can be found with this article online at http://www.current-biology.com/cgi/content/full/17/1/26/DC1/.

Acknowledgments

Work was supported by the Gatsby Charitable Foundation. We thank Keith May for help in programming the stimulus, and him, Peter Dayan, Chris Frith, Uta Frith, Sheng He, Li Jingling, and Alex Lewis for conversations and comments on our works, manuscripts, and references. Comments by the three anonymous reviewers are also much appreciated.

Received: June 22, 2006
Revised: October 12, 2006
Accepted: October 24, 2006
Published: January 8, 2007

References

1. Treisman, A.M., and Gelade, G. (1980). A feature-integration theory of attention. Cognit. Psychol. *12*, 97–136.
2. Li, Z. (2002). A saliency map in primary visual cortex. Trends Cogn. Sci. *6*, 9–16.
3. Tehovnik, E.J., Slocum, W.M., and Schiller, P.H. (2003). Saccadic eye movements evoked by microstimulation of striate cortex. Eur. J. Neurosci. *17*, 870–878.
4. Stankiewicz, B.J., Hummel, J.E., and Cooper, E.E. (1998). The role of attention in priming for left-right reflections of object images: Evidence for a dual representation of object shape. J. Exp. Psychol. *24*, 732–744.
5. Treisman, A.M., and Kanwisher, N.G. (1998). Perceiving visually presented objects: Recognition, awareness, and modularity. Curr. Opin. Neurobiol. *8*, 218–226.
6. Frith, U. (1974). A curious effect with reversed letters explained by a theory of schema. Percept. Psychophys. *16*, 113–116.
7. Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. *195*, 215–243.
8. Kahneman, D., Treisman, A., and Gibbs, B.J. (1992). The reviewing of object files: object-specific integration of information. Cognit. Psychol. *24*, 175–219.
9. Riesenhuber, M., and Poggio, T. (2003). How the visual cortex recognizes objects: The tales of the standard model. In The Visual Neurosciences, *Volume 2*, L.M. Chalupa and J.S. Werner, eds. (Cambridge, MA: MIT Press), pp. 1640–1653.
10. Tanaka, K. (2003). Inferotemporal response properties. In The Visual Neurosciences, *Volume 2*, L.M. Chalupa and J.S. Werner, eds. (Cambridge, MA: MIT Press), pp. 1151–1164.
11. Rolls, E.T. (2003). Invariant object and face recognition. In The Visual Neurosciences, *Volume 2*, L.M. Chalupa and J.S. Werner, eds. (Cambridge, MA: MIT Press), pp. 1165–1178.
12. Logothetis, N.K., Pauls, J., and Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. Curr. Biol. *5*, 552–563.
13. Humphreys, G.W., Riddoch, M.J., and Price, C.J. (1997). Top-down processes in object identification: Evidence from experimental psychology, neuropsychology and functional anatomy. Philos. Trans. R. Soc. Lond. B. Biol Sci. *352*, 1275–1282.
14. Kourtzi, Z., and Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. Science *293*, 1506–1509.
15. Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. Vision Res. *41*, 1409–1422.
16. Knierim, J.J., and Van Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J. Neurophysiol. *67*, 961–980.
17. Sillito, A.M., Grieve, K.L., Jones, H.E., Cudeiro, J., and Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. Nature *378*, 492–496.
18. Nothdurft, H.C., Gallant, J.L., and Van Essen, D.C. (1999). Response modulation by texture surround in primate area V1: Correlates of “popout” under anesthesia. Vis. Neurosci. *16*, 15–34.
19. Schiller, P.H., and Lee, K. (1991). The role of the primate extrastriate area V4 in lesion. Science *251*, 1251–1253.
20. Mazer, J.A., and Gallant, J.L. (2003). Goal-related activity in V4 during free viewing visual search: Evidence for a ventral stream visual salience map. Neuron *40*, 1241–1250.
21. Ogawa, T., and Komatsu, H. (2004). Target selection in area V4 during a multidimensional visual search task. J. Neurosci. *24*, 6371–6382.
22. Wolfe, J.M., and Bennett, S.C. (1997). Preattentive object files: Shapeless bundles of basic features. Vision Res. *37*, 25–43.
23. Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In Analysis of Visual Behavior, D.J. Ingle, M.A. Goodale, and R.J.W. Mansfield, eds. (Cambridge, MA: The MIT Press), pp. 549–586.
24. Chelazzi, L., Miller, E.K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. Nature *363*, 345–347.
25. Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J. Neurophysiol. *70*, 909–919.
26. Hoffman, J.E. (1998). Visual attention and eye movements. In Attention, H. Pashler, ed. (London: University College London Press), pp. 119–154.
27. Duncan, J., and Humphreys, G.W. (1989). Visual search and stimulus similarity. Psychol. Rev. *96*, 433–458.
28. Potter, M.C. (1976). Short-term conceptual memory for pictures. J. Exp. Psychol. [Hum. Learn.] *5*, 509–522.
29. Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. Nature *381*, 520–522.
30. Luck, S.J., Vogel, E.K., and Shapiro, K.L. (1996). Word meanings can be accessed but not reported during the attentional blink. Nature *382*, 616–618.
31. van Zoest, W., and Donk, M. (2004). Bottom-up and top-down control in visual search. Perception *33*, 927–937.
32. Gilchrist, I.D., Heywood, C.A., and Findlay, J.M. (2003). Visual sensitivity in search tasks depends on the response requirement. Spat. Vis. *16*, 277–293.
33. Fang, F., and He, S. (2005). Nat. Neurosci. *8*, 1380–1385.
34. Enns, J.T., and Di Lollo, V. (2000). What’s new in visual masking? Trends Cogn. Sci. *4*, 345–352.
35. Frazor, R.A., Albrecht, D.G., Geisler, W.S., and Crane, A.M. (2004). Visual cortex neurons of monkeys and cats: Temporal dynamics of the spatial frequency response function. J. Neurophysiol. *91*, 2607–2627.
36. Li, Z. (1992). Different retinal ganglion cells have different functional goals. Int. J. Neural Syst. *3*, 237–248.
37. Richards, J.T., and Reicher, G.M. (1978). The effect of background familiarity in visual search - An analysis of underlying factors. Percept. Psychophys. *23*, 499–505.
38. Shen, J., and Reingold, E.M. (2001). Visual search asymmetry: The influence of stimulus familiarity and low-level features. Percept. Psychophys. *63*, 464–475.