

Opposite Modulation of High- and Low-Level Visual Aftereffects by Perceptual Grouping

Dongjun He,¹ Daniel Kersten,^{4,5} and Fang Fang^{1,2,3,*}

¹Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education)

²Peking-Tsinghua Center for Life Sciences

³IDG/McGovern Institute for Brain Research

Peking University, Beijing 100871, P.R. China

⁴Department of Psychology, University of Minnesota,

Minneapolis, MN 55455, USA

⁵Department of Brain and Cognitive Engineering, Korea

University, Seoul 136-713, South Korea

(Figure 1A) with three horizontal occluders (Figure 1B). Only the tilted bars were visible to subjects. The thin diamond translated with a circular trajectory, thereby maintaining the bars at

Summary

A fundamental task of visual perception is to group visual features—sometimes spatially separated and partially occluded—into coherent, unified representations of objects. Perceptual grouping can vastly simplify the description of a visual scene and is critical for our visual system to understand the three-dimensional visual world. Numerous neurophysiological and brain imaging studies have demonstrated that neural mechanisms of perceptual grouping are characterized by the enhancement of neural responses throughout the visual processing hierarchy, from lower visual areas processing grouped features to higher visual areas representing objects and shapes from grouping [1–3]. In a series of psychophysical adaptation experiments, we made the counterintuitive observation that perceptual grouping amplified the shape aftereffect but meanwhile, reduced the tilt aftereffect and the threshold elevation aftereffect (TEAE). Furthermore, the modulation of perceptual grouping on the TEAE showed a partial interocular transfer. This finding suggests a 2-fold effect of perceptual grouping—enhancing the high-level shape representation and attenuating the low-level feature representation even at a monocular level. We propose that this effect is a functional manifestation of a predictive coding scheme [4–8] and reflects an efficient code of visual information across lower and higher visual cortical areas.

Results

We used adaptation to explore the effect of perceptual grouping on visual pattern representation in the human visual system. Adaptation is a general property of almost all neural systems. Due to its power to isolate and temporarily reduce the contribution of specific neural populations, measuring the aftereffects of adaptation has been a powerful tool of psychophysics to study the representation of various visual patterns, from low-level features to high-level shapes, objects, and faces [9–11].

In the current study, adapting stimuli were a partially occluded diamond (the diamond stimulus) and its variant (the nondiamond stimulus) (Figure 1). The diamond stimulus was constructed by masking a complete thin diamond

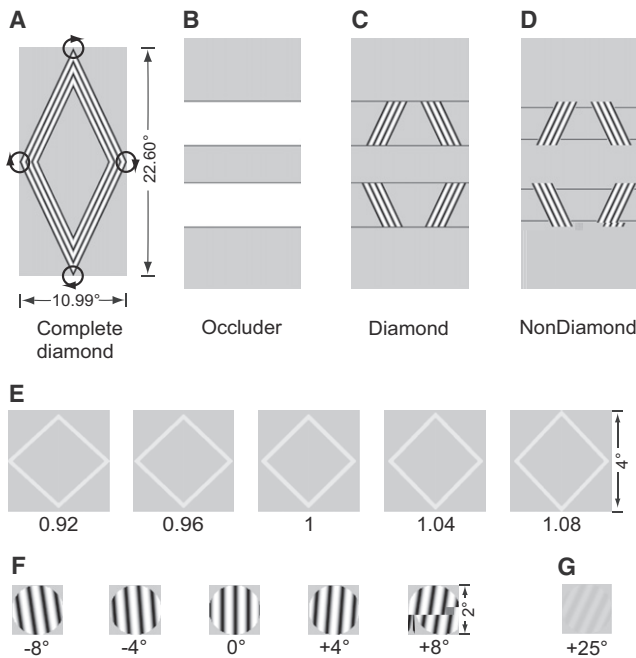


Figure 1. Visual Stimuli

(A) A complete thin diamond translated with a circular trajectory. Its direction (clockwise or counterclockwise) reversed every 5 s.
 (B) Three horizontal occluders were rendered with the background color except part of their borders—the four horizontal lines.
 (C) The diamond stimulus as adaptor was generated by masking the complete diamond with the occluders. T-junctions formed by the horizontal lines, and the visible parts of the complete diamond made subjects see a coherently translating diamond.
 (D) The nondiamond stimulus as a second adaptor was generated by displacing the horizontal lines vertically. The absence of the T-junctions broke the stimulus into four separate moving bars.
 (E) Diamond test stimuli for measuring shape aftereffect.
 (F) High-contrast grating test stimuli for measuring tilt aftereffect.
 (G) A sample low-contrast grating test stimulus for measuring threshold elevation aftereffect.

perceived normal diamond caused by adaption were taken as the magnitude of the TAE and the SAE respectively. If adaptation could generate a significant TAE and/or SAE, a vertical grating would be perceived to be left tilted and/or a normal diamond to be fat.

For three experimental conditions—adapting to the diamond stimulus, adapting to the nondiamond stimulus and baseline (without adaptation)—the perceived verticals (mean \pm SEM) were $1.84^\circ \pm 0.76^\circ$, $3.31^\circ \pm 0.90^\circ$, and $0.29^\circ \pm 0.72^\circ$, respectively. TAEs were significant after adapting to both the nondiamond stimulus ($t = 4.90$, $p < 0.01$) and the diamond stimulus ($t = 3.31$, $p < 0.05$). The TAE from the nondiamond stimulus was significantly larger than that from the diamond stimulus ($t = 6.85$, $p < 0.01$) (Figure 2B). However, SAE measurements had a distinctive pattern. The aspect ratios of the perceived normal diamonds were 0.9937 ± 0.009 , 0.9752 ± 0.01 , and 0.9743 ± 0.01 for the three conditions. A significant SAE was found after adapting to the diamond stimulus ($t = 8.21$, $p < 0.01$), but not the nondiamond stimulus ($t = 0.49$, $p > 0.05$). The difference between the two adapting stimuli was significant ($t = 5.07$, $p < 0.01$) (Figure 2C). These results demonstrate that perceptual grouping could enhance the representation of the diamond shape but attenuate the representation of the

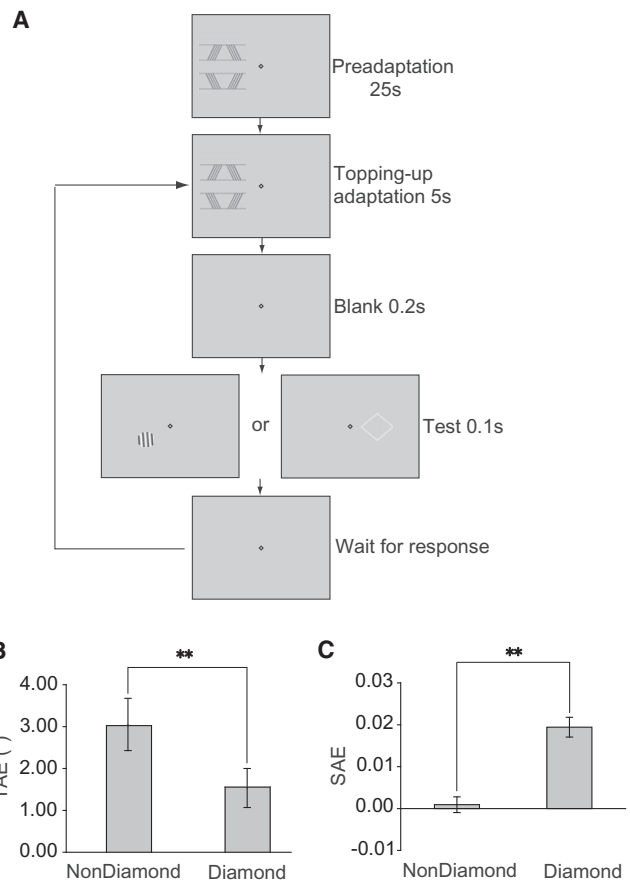


Figure 2. Procedure and Results of Experiment 1

(A) Experimental procedure. An adaptation block had only one adapting stimulus (the diamond stimulus or the nondiamond stimulus) presented in the left visual field. It began with a 25 s preadaptation. In a trial, after a 5 s topping-up adaptation and a 0.2 s blank interval, a test stimulus was presented for 0.1 s and subjects were asked to make a 2-AFC judgment. The test stimulus could be one of the five grating test stimuli presented in the left visual field for measuring TAE, and subjects needed to indicate that the grating was left or right tilted. Alternatively, the test stimulus could be one of the five diamond test stimuli presented in the right visual field for measuring SAE, and subjects judged whether the diamond was thin or fat.
 (B) TAE magnitudes from adapting to the diamond and the nondiamond stimuli.
 (C) SAE magnitudes from adapting to the diamond and the nondiamond stimuli. Asterisks indicate a statistically significant difference between two stimulus conditions (* $p < 0.05$; ** $p < 0.01$). Error bars denote 1 SEM calculated across subjects for each condition.

bar orientation. The shape adaptation should take place in high-level visual areas because the SAE was evident even when the adapting and test stimuli were presented in the left and right visual fields, respectively. A possible area is the lateral occipital area (LO) because the LO in either hemisphere is responsive to shape images presented in both the left and right visual fields [16], although it still has a contralateral preference [17].

Effect of Perceptual Grouping on Threshold Elevation Aftereffect

It could be argued that, in experiment 1, the TAE reduction by perceptual grouping is due to different spatial distributions of attention when subjects viewed the diamond and the

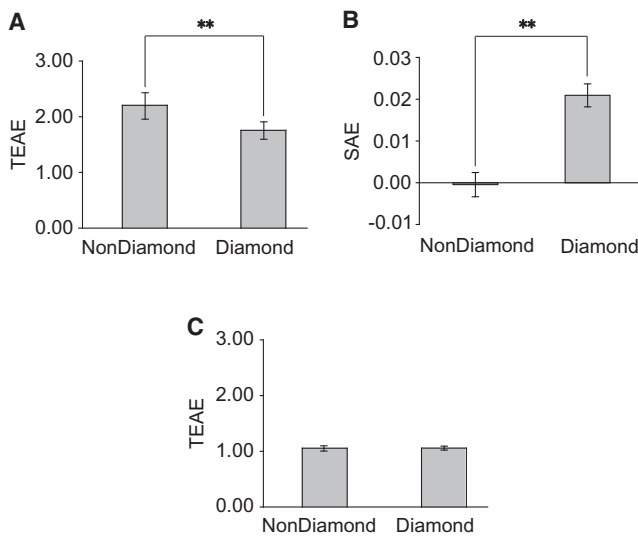


Figure 3. Results of Experiment 2

(A) TEAE magnitudes from adapting to the diamond and the nondiamond stimuli when the adapted and test orientations were identical.

(B) SAE magnitudes from adapting to the diamond and the nondiamond stimuli.

(C) TEAE magnitudes from adapting to the diamond and the nondiamond stimuli when the test orientation was orthogonal to the adapted orientation. Asterisks indicate a statistically significant difference between two stimulus conditions (* $p < 0.05$; ** $p < 0.01$). Error bars denote 1 SEM calculated across subjects for each condition.

nondiamond stimuli. For the nondiamond stimulus, subjects' spatial attention might be more focused on the four moving bars. For the diamond stimulus, attention might even spread to the occluders, the hidden corners, and the area bound by the four bars because they are intrinsically related to each other and construct a representational entity [18]. To confirm that the representation attenuation of the bar orientation is due to perceptual grouping rather than a pure attentional effect, we performed experiment 2 with five other subjects to measure the effect of perceptual grouping on contrast threshold elevation aftereffect (TEAE). Although TAE could be modulated by attention, threshold elevation aftereffect (TEAE) has been demonstrated to be independent of attention, especially when adapting stimuli have a high contrast [19–21].

Experiment 2 measured the TEAE and the SAE from adapting to the diamond and the nondiamond stimulus. For the TEAE measurement, we used a temporal 2-AFC QUEST staircase procedure (82% correct) [22] to measure subjects' contrast detection thresholds with and without adaptation. The ratio of the threshold with adaptation to that without adaptation was taken as the TEAE magnitude. The orientation of test stimuli was identical to that of the adapting stimulus (i.e., the lower right bar) (Figure 1G). We found a significant TEAE after adapting to both the nondiamond stimulus ($t = 9.31$, $p < 0.01$) and the diamond stimulus ($t = 11.21$, $p < 0.01$). The TEAE from the nondiamond stimulus was significantly larger than that from the diamond stimulus ($t = 5.02$, $p < 0.01$) (Figure 3A). The SAE measurement was performed in the same way as that in experiment 1 and it replicated the finding (Figure 3B). In addition, we also measured the TEAE when the orientation of test stimuli was orthogonal to that of the adapting stimulus. No significant effect was found, which

confirmed that this local adaptation was orientation-specific (Figure 3C). These results provide further evidence that perceptual grouping could attenuate the representation of the bar orientation.

Interocular Transfer of Perceptual Grouping Effect

Because TAE and TEAE are believed to be generated in early visual cortex, even as early as in V1 (especially for TEAE) [20, 23], the results above suggest that perceptual grouping could attenuate the representation of the bar orientation in early visual cortical areas. To further examine whether some monocular mechanism in V1 contributes to the perceptual grouping effect in experiment 2 (i.e., the TEAE reduction), we performed experiment 3 to measure the amount of the interocular transfer of the TEAE reduction. The TEAE measurement was similar to that in experiment 2 except that adapting and test stimuli were presented in either the same eye or different eye. Subjects viewed the stimuli through a mirror stereoscope. Adapting stimuli were presented in either the left or the right eye (Figure 4A). Test stimuli were always presented in the left eye (Figure 4B). We quantified the effect of perceptual grouping on TEAE with an index ($1 - \frac{\text{TEAE}_{\text{different eye}}}{\text{TEAE}_{\text{same eye}}}$). A large index means a strong effect of perceptual grouping. Interocular transfer refers to the relative size of the grouping effect when the adapting and test stimuli are presented to different eyes compared to when presented to the same eye. A 100% transfer means that the different-eye and same-eye effects are of the same magnitude and indicates a purely binocular process. A small transfer suggests that the mechanism involved is primarily monocular. Intermediate levels of transfer are best explained in terms of a mixture of monocular and binocular mechanisms [24] (but see also [25]).

Regardless of whether the adapting and test stimuli were presented in the same eye or different eyes, we found a significant TEAE after adapting to both the nondiamond stimulus (same eye: $t = 11.46$, $p < 0.01$; different eye: $t = 28.60$, $p < 0.01$) and the diamond stimulus (same eye: $t = 14.43$, $p < 0.01$; different eye: $t = 29.46$, $p < 0.01$). The TEAE from the nondiamond stimulus was significantly larger than that from the diamond stimulus (same eye: $t = 6.46$, $p < 0.01$; different eye: $t = 8.01$, $p < 0.01$) (Figures 4C and 4D). These results are consistent with the finding in experiment 2. An interesting finding in this experiment is that the effect of perceptual grouping on TEAE in the same eye condition was significantly larger than that in the different eye condition ($t = 7.79$, $p < 0.01$). The indices of the perceptual grouping effect for the two conditions are 0.22 and 0.11, respectively (Figure 4E). Thus, the interocular transfer of the effect was 48.4%, suggesting a mixture of monocular and binocular mechanisms underlying the grouping effect. A caveat should be noted that the extent of eye specificity inferred from the above analysis might be overestimated because of a higher level of measurement noise in the same eye condition, as indicated by the larger error bars in Figure 4C than those in Figure 4D.

Finally, we carried out experiments demonstrating that the observed effects in experiments 1–3 were not due to the physical difference between the diamond and the nondiamond stimuli. Their difference was the tiny position changes of the four horizontal lines. We had subjects adapt to four horizontal lines whose positions were identical to those in the diamond stimulus or the nondiamond stimulus. No detectable TAE, TEAE, or SAE was observed after adaptation.

Discussion

Our experiments provide clear evidence that perceptual grouping could magnify the high-level SAE but reduce the low-level TAE and TEAE. These results demonstrate that a functional role of perceptual grouping is enhancing the high-level shape representation and meanwhile weakening the representation of the constituent elements (i.e., bar orientations) of the shape. Moreover, the effect of perceptual grouping on the TEAE showed a partial interocular transfer—it was significantly reduced when the adapting and test stimuli are presented to different eyes compared to when presented

in inverse activity patterns. Predictive coding models [4–8], for example, are one class of models that suggest that feedback may operate to reduce activity in lower areas. These models posit that higher areas are actively attempting to “explain” activity patterns in lower areas via feedback projections. Because most predictive coding models include a subtractive comparison between the hypotheses formed in higher areas and the incoming sensory input represented in lower areas, the overall effect of feedback may be to reduce activity in lower areas. Specifically, reduced activity in lower visual areas would occur whenever the predictions of higher level areas match incoming sensory information. In the case of our stimuli, when high visual areas (e.g., LO) maintain a representation of a grouped shape, this “expectation” or “understanding” of the image features is sent back to lower visual areas (e.g., V1) and removed, resulting in less activity. When higher areas are unable to form such an understanding (i.e., when the bars are perceived as ungrouped), these feedback processes are not occurring and there is consequently more activity in lower areas. Predictive coding models have strong intuitive appeal— why bother signaling what you already know [32]? The reduced activity that would result from such a process would also have substantial biological benefits. There are clear efficiency constraints placed on the visual system—both because of inherent capacity limitations in neural pathways and because spikes are metabolically expensive [33]. The visual system would do well to use a representational strategy that maximizes biological efficiency by utilizing a code that minimizes spike rate.

Recent fMRI and magnetoencephalography (MEG) studies have provided evidence for predictive coding models [34–37]. We also performed fMRI experiments to test the models. We observed a BOLD signal increase in the LO and a concurrent signal decrease in V1 when visual elements were assembled into a coherent shape [14, 15]. However, it should be pointed out that, due to the complicated nature of BOLD signal and the limit of its spatial resolution [38], BOLD signal reductions in lower areas cannot be unequivocally explained as a decrease in neural activity representing low-level elements [32]. The reductions may be a manifestation of representation sharpening or noise removal [39]. They could also be attributed to other factors, including changes in visual stimulus, perceived context, and attentional state. More critically, behavioral significance of predictive coding has rarely been verified. By showing that shape perception from perceptual grouping affects not only high-level vision, but also low-level vision, the current adaptation study provides the first piece of behavioral evidence for a predictive coding scheme.

Our study suggests that feedback from higher visual areas serves to reduce activity in lower visual areas during perceptual grouping. The feedback could even penetrate back to monocular neurons in V1. It should be noted that a major challenge for the predictive coding view is how a higher visual area predicts the precise metrics of a stimulus. A dominant functional interpretation of the feedforward ventral pathway is increased selectivity at the expense of insensitivity to variables such as translation, illumination, and scale; but if information about position and size is gradually lost, then how could a feedback signal be spatially precise? One possibility is that transformation information might be retained in the visual processing hierarchy, as suggested by a recent study [40]. The predictive coding view is an alternative proposal to the conventional wisdom that favors enhancement and attention in the conceptualization of the role of feedback

in visual processing. Understanding how the feedback is implemented in the visual cortex will be a scientific challenge in the future.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures and can be found with this article online at [doi:10.1016/j.cub.2012.04.026](https://doi.org/10.1016/j.cub.2012.04.026).

Acknowledgments

This work was supported by the Ministry of Science and Technology of China (2011CBA00405 and 2010CB833903), the National Natural Science Foundation of China (Project 30925014, 30870762, and 90920012), and the Fundamental Research Funds for the Central Universities. D.K. was partially supported by the World Class University program funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008) and by the National Institutes of Health grant R01 EY015261.

Received: February 11, 2012

Revised: March 28, 2012

Accepted: April 13, 2012

Published online: May 10, 2012

References

1. Kourtzi, Z., Tolias, A.S., Altmann, C.F., Augath, M., and Logothetis, N.K. (2003). Integration of local features into global shapes: monkey and human fMRI studies. *Neuron* 39, 333–346.
2. Roelfsema, P.R. (2006). Cortical algorithms for perceptual grouping. *Annu. Rev. Neurosci.* 29, 203–227.
3. Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron* 54, 677–696.
4. Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* 66, 241–251.
5. Barlow, H.B. (1994). What is the computational goal of the neocortex? In *Large-Scale Neuronal Theories of the Brain*, C. Koch and J.L. Davis, eds. (Cambridge, MA: MIT), pp. 1–22.
6. Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87.
7. Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271–304.
8. Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138.
9. Blakemore, C., and Campbell, F.W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol.* 203, 237–260.
10. Webster, M.A., Kaping, D., Mizokami, Y., and Duhamel, P. (2004). Adaptation to natural facial categories. *Nature* 428, 557–561.
11. Fang, F., and He, S. (2005). Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. *Neuron* 47, 793–800.
12. Lorenceau, J., and Alais, D. (2001). Form constraints in motion binding. *Nat. Neurosci.* 4, 745–751.
13. McDermott, J., Weiss, Y., and Adelson, E.H. (2001). Beyond junctions: nonlocal form constraints on motion interpretation. *Perception* 30, 905–923.
14. Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., and Woods, D.L. (2002). Shape perception reduces activity in human primary visual cortex. *Proc. Natl. Acad. Sci. USA* 99, 15164–15169.
15. Fang, F., Kersten, D., and Murray, S.O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *J. Vis.* 8, 1–9.
16. Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res.* 41, 1409–1422.
17. Niemeier, M., Goltz, H.C., Kuchinad, A., Tweed, D.B., and Vilis, T. (2005). A contralateral preference in the lateral occipital area: sensory and attentional mechanisms. *Cereb. Cortex* 15, 325–331.
18. Naber, M., Carlson, T.A., Verstraten, F.A.J., and Einhauser, W. (2011). Perceptual benefit of objecthood. *J. Vis.* 11, 1–9.

19. Festman, Y., and Ahissar, M. (2004). Attentional states and the degree of visual adaptation to gratings. *Neural Netw.* *1* , 849–860.
20. Blake, R., Tadin, D., Sobel, K.V., Raissian, T.A., and Chong, S.C. (2006). Strength of early visual adaptation depends on visual awareness. *Proc. Natl. Acad. Sci. USA* *103*, 4783–4788.