Face perception inherits low-level binocular adaptation

Keith A. May

Department of Psychology, University of Essex,
Colchester, UK

UCL Department of Computer Science,
University College London, London, UK
Present addresses: Center for Integrative Neuroscience,
University of Tübingen, Tübingen, Germany; and
Department of Sensory and Sensorimotor Systems,
Max Planck Institute for Biological Cybernetics,
Tübingen, Germany

Li Zhaoping

In previous work (May & Zhaoping, 2016; May, Zhaoping, & Hibbard, 2012), we have provided evidence that the visual system efficiently encodes binocular information using separately adaptable binocular summation and differencing channels. In that work, binocular test stimuli delivered different grating patterns to the two binocular channels; selective adaptation of one of the binocular channels made participants more likely to see the other channel's grating pattern. In the current study, we extend this paradigm to face perception. Our test stimuli delivered different face images to the two binocular channels, and we found that selective adaptation of one binocular channel biased the observer to perceive the other channel's face image. We show that the perceived identity, gender, emotional expression, or direction of 3-D rotation of a facial test image can be influenced by pre-exposure to binocular random-noise patterns that contain no meaningful spatial structure. Our results provide compelling evidence that face-processing mechanisms can inherit adaptation from low-level sites. Our adaptation paradigm targets the low-level mechanisms in such a way that any response bias or inadvertent adaptation of high-level mechanisms selective for face categories would reduce, rather than produce, the measured effects of adaptation.

Introduction

Li and Atick (1994) proposed that the two eyes' signals are coded efficiently in the brain using binocular summation and differencing channels very early in the processing stream (see Figure 1). According to the theory, each of these channels can be independently desensitized temporarily by strong stimulation in that channel, leading to adaptation effects.

We have previously reported psychophysical evidence for these selectively adaptable summation and differencing channels (May & Zhaoping, 2016; May, Zhaoping, & Hibbard, 2012). In the paradigm that we developed, one eye's stimulus $(A + B)$ is the sum of two stimuli, A and B, while the other eye's stimulus $(A - B)$ is the difference between A and B. In the binocular summation channel, the visual input is the sum of the input to the two eyes, $(A + B) + (A - B)$, so the B components cancel out, leaving A; in the binocular differencing channel, the A components cancel out, leaving B. Whether the participant perceives stimulus A or B depends on the relative sensitivity of the two binocular channels. We found that by selectively adapting one or the other of the binocular channels using binocular random-noise stimuli, we could bias perception toward A or B. In our research so far, the component stimuli (A and B) have been simple grating patterns with different motion directions (May et al., 2012) or orientations (May & Zhaoping, 2016). In the study reported here, we extend this paradigm to human faces, showing that perception of identity, gender, emotional expression, or direction of 3-d head rotation can be influenced by pre-exposure to binocular random-noise patterns.

The extension of this paradigm to faces is worthwhile because it allows us to examine whether or not low-level adaptation effects believed to occur in the primary visual cortex (V1) can be inherited by the high-level face-processing areas of the cortex. Although one might expect adaptation at low levels of processing to affect processing at higher levels, another possibility outlined by Xu, Dayan, Lipkin, and Qian (2008) is "that adaptive changes to a code (at lower levels) are precisely tracked by higher levels, eliminating corruption from low-level adaptation" (p. 3375). They state

A   Binocular summation channel

αH

$+$

Eye 1

$(\alpha H + \beta S)/2$

Eye 2

$(\alpha H - \beta S)/2$

$-$

Binocular differencing channel

βS

B   Binocular summation channel

αS

$+$

Eye 1

$(\alpha S + \beta H)/2$

Eye 2

$(\alpha S - \beta H)/2$

$-$

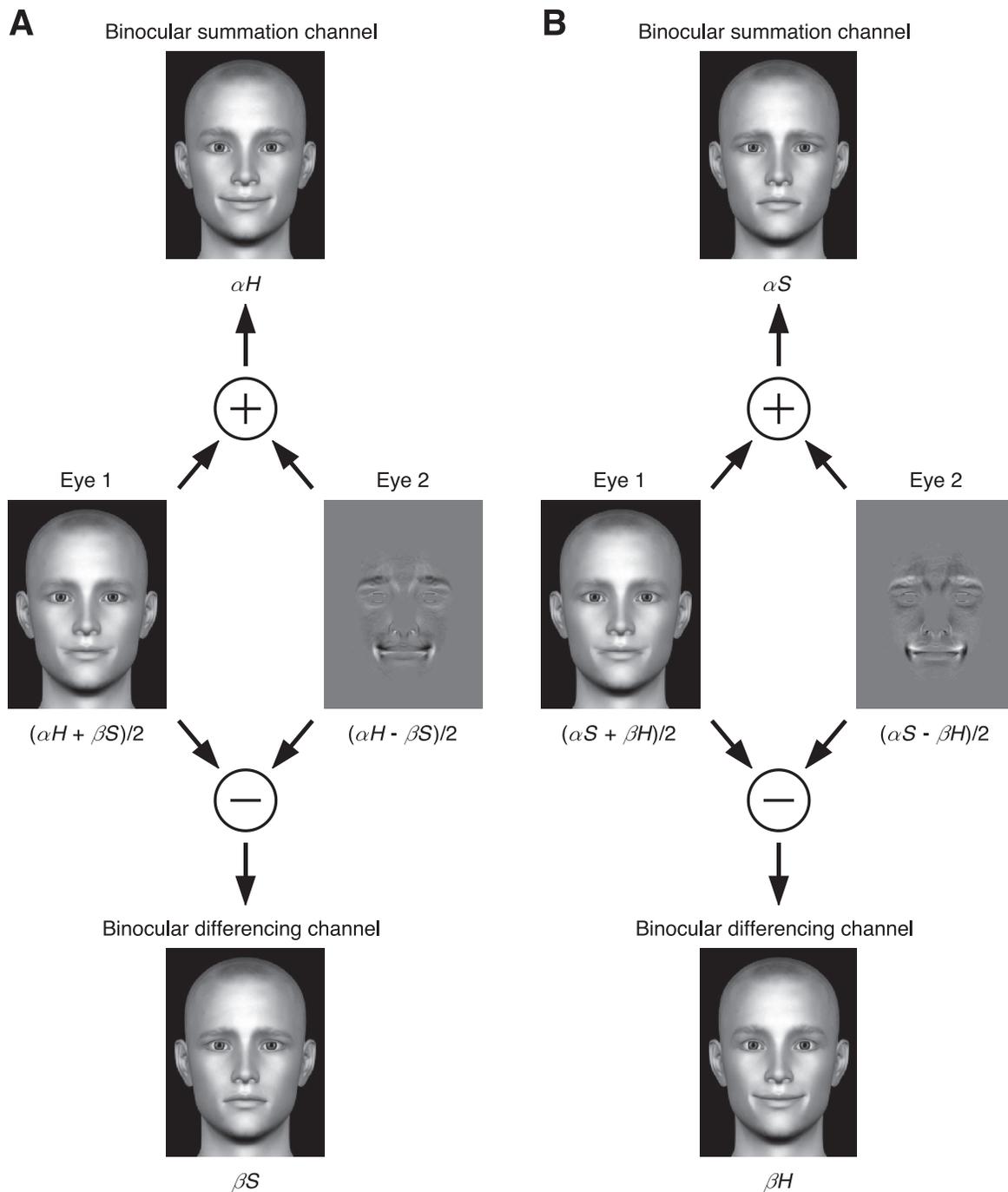Binocular differencing channel

βH

Figure 1. Li and Atick's (1994) theory of binocular encoding, with happy/sad facial test images. (A) One eye views a weighted sum of the two original images *H* (happy) and *S* (sad), and the other eye views a weighted difference of the two original images. The binocular summation channel receives a happy-face image with contrast α, and the binocular differencing channel receives a sad-face image with contrast β. (B) The contrast polarity and weightings of the composite portraits are adjusted so that the emotional expressions are swapped between the channels.

that "this issue is critical for understanding cortical coding and computation" (p. 3374), and this motivates the extension of our adaptation paradigm to face perception.

There is accumulating physiological evidence that at early stages of visual processing, adaptation at one processing stage is inherited by subsequent stages. Solomon, Pierce, Dhruv, and Lennie (2004) have reported contrast adaptation in M-cells of the macaque lateral geniculate nucleus (LGN) that is well accounted for by inheritance of adaptation from the retinal ganglion cells driving these LGN cells. Dhruv, Tailby,

Sokol, and Lennie (2011) have presented evidence that multiple sites of adaptation contribute to the adaptation observed in macaque V1 neurons. They argue that adaptation due to high-temporal-frequency adaptation stimuli of any orientation is inherited from M-cells of the LGN; adaptation due to low-temporal-frequency adaptors oriented orthogonally to the V1 neuron's preferred orientation is inherited from neurons in the input layer of V1; and adaptation due to low-temporal-frequency adaptors at the V1 neuron's preferred orientation originates largely in the V1 neuron itself. Working on mice, Dhruv and Carandini (2014) have shown that the qualitatively different effects of adaptation in LGN and V1 (change of response gain in LGN, compared with mainly a change of preferred position in V1) can be accounted for by a model in which the adaptation of LGN cells is inherited by V1 cells via connections that are unaffected by adaptation. Further downstream, Kohn and Movshon (2003) have shown that the reduced contrast sensitivity of macaque MT neurons after motion adaptation shows spatial specificity that cannot be explained by adaptation processes intrinsic to MT neurons: When adaptation and test stimuli were in different parts of the MT neuron's receptive field, there was almost no change of contrast sensitivity, suggesting that this aspect of adaptation is entirely inherited from earlier stages, most likely V1. On the other hand, they (Kohn & Movshon, 2004) and Priebe, Churchland, and Lisberger (2002) have reported characteristics of MT adaptation that appear to arise within MT itself. Finally, using spatial specificity as an index of the origin of fMRI adaptation, Larsson and Harrison (2015) have found evidence that fMRI adaptation in most human extrastriate visual areas (apart from MT) originates in V1.

The physiological work showing inheritance of adaptation has focused on early visual processing stages that analyze relatively simple features of the image. It is therefore interesting to ask whether inheritance of adaptation continues to very high-level cortical areas that process complex perceptual categories such as faces. Face adaptation is a very well established phenomenon, whereby prolonged viewing of face images can influence the perception of subsequently presented faces (Webster & MacLeod, 2011; Webster & MacLin, 1999). For example, a face looks more masculine or more feminine after pre-exposure to, respectively, female or male faces (Webster, Kaping, Mizokami, & Duhamel, 2004).

Dickinson and colleagues argue that face-adaptation effects are largely inherited from adaptation at lower-level sites. They have reported several face-adaptation effects that can be fully explained by a tilt aftereffect generated in low-level mechanisms (Dickinson, Almeida, Bell, & Badcock, 2010; Dickinson & Badcock, 2013;

Dickinson, Mighall, Almeida, Bell, & Badcock, 2012). They argued that prolonged viewing of faces gives rise to orientation-selective adaptation at low-level sites; when a test face is presented after adaptation, the local low-level adaptation causes a local tilt aftereffect at each point in the visual field, with different tilt aftereffects at different locations. The *tilt aftereffect field* describes the tilt aftereffect generated at each point in the visual field. These tilt aftereffects are said to distort perception of the test face.

Other researchers argue that face-adaptation effects are at least partly caused by adaptation of high-level brain mechanisms involved in face processing. For example, Afraz and Cavanagh (2009) argue that face aftereffects show transfer across retinal position, orientation, and size that far exceed the tuning widths of low-level mechanisms early in the processing stream; they argue that their results are consistent with the site of adaptation being the high-level mechanisms that process faces. Similar arguments have been made by many other researchers (e.g., Hills & Lewis, 2012; Hole, 2011; Leopold, O'Toole, Vetter, & Blanz, 2001; Vakli, Németh, Zimmer, Schweinberger, & Kovács, 2012; Watson & Clifford, 2003; Yamashita, Hardy, De Valois, & Webster, 2005; Zhao & Chubb, 2001). On the other hand, Dickinson and Badcock (2013) argue that the transfer across retinal position in these studies may have been mediated by inadvertent eye movements, while transfer across changes in size and orientation could still have been mediated by low-level processes.

Stronger evidence in favor of a high-level site of face adaptation is that, although adaptation to a distorted face distorts the perception of a normal face, adaptation to a normal face does not distort the perception of a distorted face (figure 4 of Webster & MacLin, 1999). It is difficult to see how this asymmetry could be accommodated by a low-level adaptation mechanism: The tilt aftereffect field of Dickinson and colleagues is an odd-symmetric function of the difference in local orientation between adaptation and test stimuli at each point (figure 3A of Dickinson et al., 2010), so swapping the adaptation and test images should change only the sign of the local tilt aftereffect, not its magnitude.

Xu et al. (2008) have shown that adaptation to a simple inverted-U curve can have an effect similar to adapting to a sad face, making subsequently presented face images look happier. They argue that this supports the hypothesis that low-level adaptation can affect face perception. However, face-selective neurons in higher visual areas can be stimulated by isolated facial features such as a mouth (Perrett, Rolls, & Caan, 1982), so these simple geometric adaptors could have induced adaptation of high-level mechanisms selective for sad faces. Later work has appeared to favor the high-level mechanisms as the locus of this adaptation effect: Xu, Liu, Dayan, and Qian (2012) have shown that
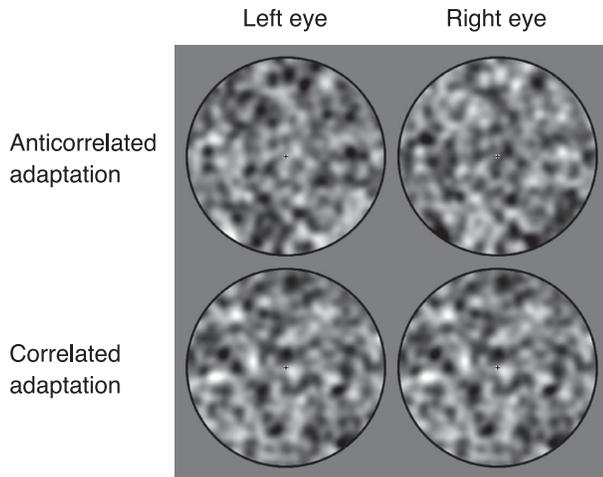
Left eye   Right eye

Anticorrelated
adaptation

Correlated
adaptation

Figure 2. Random-noise patterns used for adaptation. Each pattern was low-pass Gaussian filtered noise, surrounded by a black ring. In the "correlated adaptation" condition (lower pair), each eye received the same noise pattern; in the "anticorrelated adaptation" condition, the contrast of the noise pattern was reversed between the eyes.

crowding the geometric curve adaptors with flanking curves reduced low-level adaptation effects (curvature adaptation) but caused no significant reduction of facial-expression adaptation. This suggests that this face-adaptation effect may not be inherited from low-level mechanisms: If it were inherited, then the reduction in low-level adaptation strength due to crowding should also be inherited.

Despite the evidence that face-adaptation effects can be generated in the face-processing regions themselves, we agree with Dickinson and colleagues that it is highly plausible that face-adaptation effects can be inherited from lower levels of processing. However, it is difficult to establish this for certain because in most cases the adaptation stimuli could conceivably have selectively adapted the relevant face-processing mechanisms. Some authors have argued that the reduced strength of the face aftereffect when test and adaptation images differ in low-level visual properties reveals a contribution of low-level mechanisms (e.g., Hills & Lewis, 2012); however, as noted previously (Zhao & Chubb, 2001), some face-selective neurons in temporal cortex do show stimulus specificity (Rolls & Baylis, 1986), and so it cannot be ruled out that any incomplete transfer of the aftereffect is due to adaptation of stimulus-specific high-level face-processing mechanisms.

To establish the contribution of low-level mechanisms psychophysically, we needed to demonstrate face-adaptation effects that could not be mediated by high-level mechanisms. To do this, we took the binocular adaptation paradigm developed in our earlier work (May & Zhaoping, 2016; May et al., 2012) and extended it to judgments of face identity, gender (male/

female), emotional expression (happy/sad), and 3-D head rotation (left/right).

The images presented to the two eyes in Figure 1A are examples of the test patterns that we used for the happy/sad judgment. Each eye's image was a composite portrait (Galton, 1878) made by combining superimposed images of happy and sad faces. One eye's composite portrait was the sum of the original images (Happy + Sad); the other eye's composite portrait was the difference (Happy − Sad). In the binocular summation channel, the Sad components cancel out, leaving Happy; in the differencing channel, the Happy components cancel out, leaving Sad. Selective adaptation (i.e., desensitization) of the summation channel should make the observer more likely to perceive the sad face, whereas selective adaptation of the differencing channel should make the observer more likely to perceive the happy face. Figure 1B is similar to Figure 1A, except that the contrast polarity of the (Happy − Sad) composite portrait is reversed, causing the identities to be swapped between the channels.

Desensitization of the binocular channels was achieved by pre-exposure to random-noise patterns (see Figure 2) that were either binocularly correlated (same image in each eye) or binocularly anticorrelated (each eye saw the photographic negative of the other eye's image). The correlated noise selectively stimulates (and therefore desensitizes) the binocular summation channel, while the anticorrelated noise selectively desensitizes the binocular differencing channel. We refer to these conditions as, respectively, correlated adaptation and anticorrelated adaptation.

Within each block of trials, the participant was required to make a single type of face judgment (identity, gender, emotion, or head direction). For the emotion judgment, on each trial the participant was presented with a pair of binocular test images like those in Figure 1. On a random half of the trials, the facial test images were designed so that the summation channel would see the happy face and the differencing channel would see the sad face (Figure 1A); on the other half, the faces were swapped between the channels (Figure 1B). The participant pressed a left key to indicate happy and a right key to indicate sad. We measured the percentage of responses consistent with the summation channel's image.

Any effects of our binocular adaptation paradigm on perceived facial category must be inherited from lower levels, and could not possibly be explained by adaptation of mechanisms selective for face categories. First, since our random-noise adaptation patterns had no meaningful structure and changed to a different random pattern every half second, it is very unlikely that they would have selectively adapted high-level face-processing mechanisms. Second, even if our adaptation conditions had inadvertently differed in the

extent to which they adapted high-level mechanisms selective for face categories, this would have reduced, rather than produced, any adaptation effect. To understand this, consider the happy/sad judgment, and imagine that our adaptation patterns had selectively desensitized mechanisms tuned to the happy face; this would have biased the participant to perceive the sad face in our composite images. On half the trials, the participant's perception would have matched the summation channel's image, and on the other half of trials it would have matched the differencing channel's image, so the participant would have been at chance (50%) on our behavioral measure (percent of trials agreeing with the summation channel). Thus, any bias toward either face category (either a perceptual bias caused by high-level face-category adaptation or a cognitive response bias to respond preferentially to one of the face categories) would have reduced any measured aftereffect and could not have caused it. A response bias toward one of the response options usually has behavioral effects that are indistinguishable from a genuine perceptual aftereffect (Morgan, Dillenburger, Raphael, & Solomon, 2012). In our experiment, any response bias or high-level adaptation *in either direction* would have reduced the size of the measured adaptation effect; this allowed us to disentangle the effects of low-level adaptation from the effects of high-level adaptation to an extent that is not normally possible: Our binocular adaptation paradigm can bias the perception of face category only if the face-processing mechanisms are inheriting adaptation from lower levels in the visual-processing stream.

Methods

Participants

Five male and two female participants took part, aged between 25 and 49 years with normal or corrected-to-normal vision. All were experienced psychophysical observers, but only author KAM (squares in Figure 3) was aware of the purposes of the experiment.

Pilot data collected from KAM indicated that our face-adaptation effect was quite weak, so we deliberately selected participants who had shown a strong effect of binocular adaptation in a previous study in which the test stimuli were low-level grating patterns, not face stimuli (May & Zhaoping, 2016). In that previous study, all participants had shown an effect of adaptation in the predicted direction, but some showed substantially stronger effects than others. The fact that all participants in the previous study showed an effect in the predicted direction suggests that our current results would generalize to the population as a whole;

however, for some participants it could take a very large number of trials to obtain statistically robust results, compared with the more modest number of trials that we required with our specially selected participants.

Test stimuli

The happy, sad, left-rotated, and right-rotated images were created using a 3-D model in Daz Studio (Version 4.8; Daz Productions, Inc., Salt Lake City, UT). For most participants, the rotated heads were rotated ±2° from straight ahead in 3-D space about their vertical axis; two participants (diamonds and left-pointing triangles in Figure 3) found these too difficult to discriminate, so for these two participants we used images in which the heads were rotated ±2.5°. The male and female images were exaggerated male and female composite face images taken from Perrett et al. (1998, figure 2a, 2d). Images of actors Brad Pitt and Matt Damon and the British TV double act Ant and Dec were found on the Internet using a Google Image search. We used a single publicity image containing both Ant and Dec to maximize similarity in the photographic conditions for these two face images; however, this resulted in the two individuals being slightly rotated in opposite directions, so that their features did not match up. To fix this problem, we left/right mirror-reversed the image of Ant.

These images were initially obtained as standard RGB image files with RGB integer values between 0 and 255. Digital photographic images produced by cameras and graphics software are encoded with gamma correction. This correction applies a compressive transducer (close to a power function with exponent 1/2.2) to the luminance values in order to compensate for the screen nonlinearity (which typically approximates a power function with exponent 2.2). Therefore, before processing the images for use in the experiment, we "uncorrected" them so that the image values were approximately linear functions of intended luminance. This linearization was carried out by dividing the RGB values by 255 (to give values between 0 and 1) and then raising these values to the power of 2.2. These linearized RGB values were then converted to grayscale using the standard formula (Burger & Burge, 2009) for finding the luminance Y of an sRGB image from the linearized values:

$$Y = 0.2126R + 0.7152G + 0.0722B. \quad (1)$$

These grayscale values fall between 0 and 1, so they were linearly scaled to produce a local contrast signal $C = 2Y - 1$ that fell between -1 and 1, in order to perform the image processing described in the following.
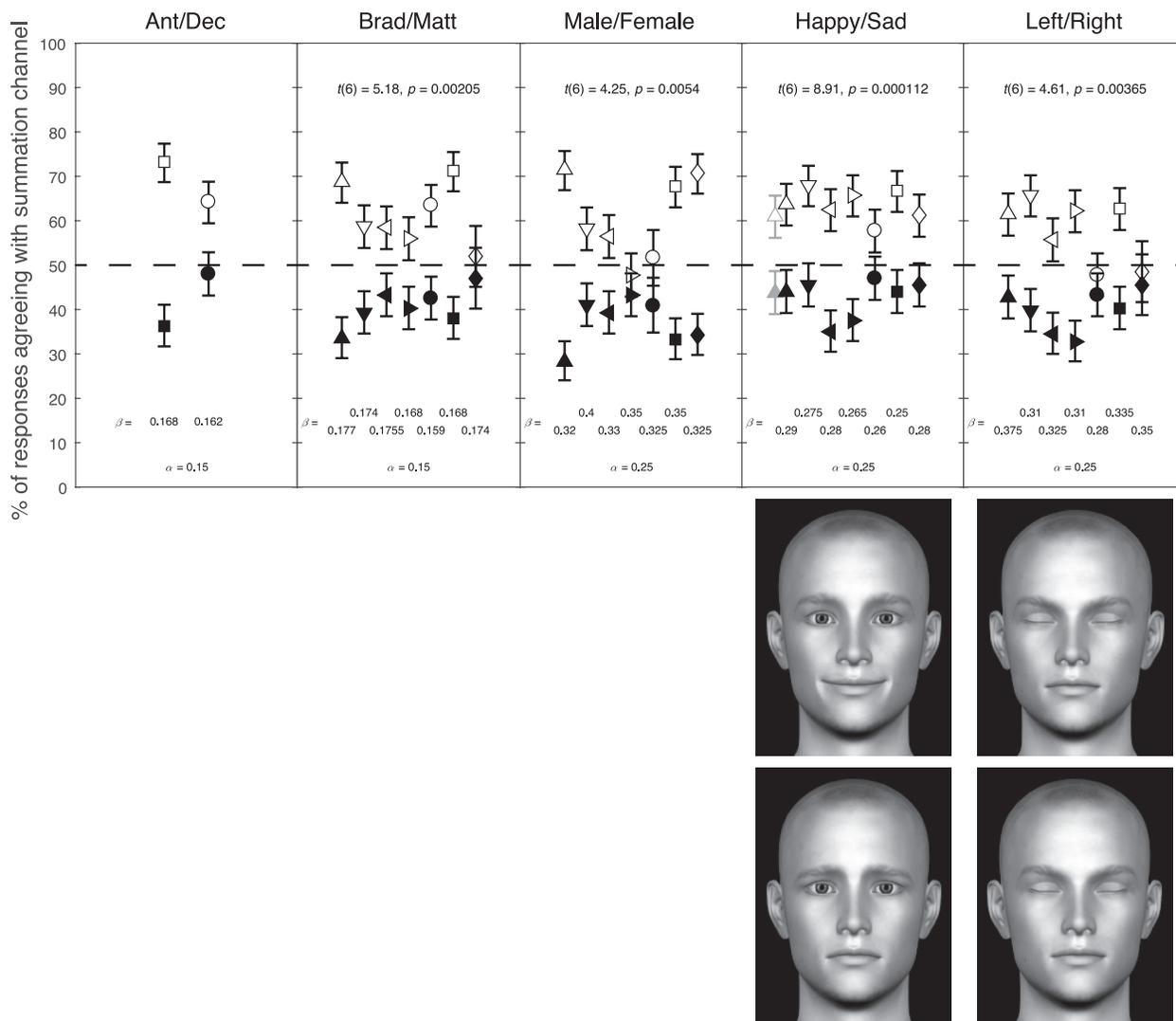
Figure 3. Results. Each column shows the results (top) from one face-category judgment. The two rightmost columns show the pair of original face images (bottom) from which the composite portraits for that category judgment were made. The celebrity images and male/female composite images are omitted for legal/copyright reasons. Different participants' data are plotted with different symbol shapes. Filled and open symbols plot data for the correlated and anticorrelated adaptation conditions, respectively. For the leftmost participant on the happy/sad judgment (upward-pointing triangles), we plot the results of a full replication of this condition in gray. For a given face pair, the summation-channel contrast α was the same for all participants, but the differencing-channel contrast β was tailored to each participant to balance the data about the 50% point. Each participant's β for each face judgment is given below the plotted points for that participant/judgment. Values of α are given below β. Error bars indicate 95% Bayes credible intervals (Nicholson, 1985). At the top of each column (except Ant/Dec) we report the results of a repeated-measures two-tailed t test of the difference between correlated and anticorrelated adaptation; the t test for happy/sad did not include the replication plotted in gray.

The grayscale images were then resized, rotated, and cropped to the following sizes (width × height in pixels): 175 × 257 (Ant/Dec), 175 × 207 (Brad/Matt), 177 × 230 (male/female), and 177 × 234 (happy/sad and left/right).

Let P be the image that corresponds to one response option (e.g., Ant, Brad, male, happy, or rotated left) and let Q be the image that corresponds to the other response option (e.g., Dec, Matt, female, sad, or rotated right). Then P and Q are 2-D matrices holding the local contrast values C (already defined) of the two images. In what follows, α and β are scalar multipliers that control the global contrasts of these images.

On half the trials, one eye received image I_1 and the other eye received image I_2:

$$I_1 = (\alpha P + \beta Q)/2 \quad (2)$$

$$I_2 = (\alpha P - \beta Q)/2. \quad (3)$$

Each image was presented to the left and right eyes equally often. Note that images I_1 and I_2 are expressed as local contrast signals, which fall between -1 and 1 about a background of 0. To convert these contrast signals I to luminance L to display them on the screen, we used the following formula:

$$L = L_0(1 + I), \quad (4)$$

where L_0 is the background luminance (54 cd/m^2). Thus, a local contrast value of -1 is mapped to zero luminance, and a local contrast value of 0 is mapped to the background luminance. The reason for describing the images in terms of local contrast is that the signal projected from the retina is essentially a contrast signal rather than a luminance signal (Troy & Enroth-Cugell, 1993), so we need to consider the contrast signals I_1 and I_2 when determining how the images will be processed in early visual processing. The responses of our putative summation (S_+) and differencing (S_-) channels would be

$$S_+ = I_1 + I_2 = \alpha P \quad (5)$$

$$S_- = I_1 - I_2 = \beta Q. \quad (6)$$

On the other half of the trials, images I_1 and I_2 were defined as follows, again with each image being presented to the left or right eye equally often:

$$I_1 = (\alpha Q + \beta P)/2 \quad (7)$$

$$I_2 = (\alpha Q - \beta P)/2. \quad (8)$$

The responses of the summation (S_+) and differencing (S_-) channels on these trials would be

$$S_+ = \alpha Q \quad (9)$$

$$S_- = \beta P. \quad (10)$$

Adaptation stimuli

The adaptation stimuli were isotropic Gaussian low-pass filtered noise (see Figure 2). Noise images with size 350×350 pixels were created in the Fourier domain with an amplitude that was a Gaussian function of spatial frequency with a standard deviation of 0.02 c/pixel and zero mean. The amplitude of the zero-frequency component was set to zero. The phase of each Fourier component was random, with the restriction that equal-frequency components on opposite sides of the origin of Fourier space had phase values with the same magnitude but opposite sign; this created a complex conjugate relationship between corresponding positive and negative frequency components, causing the imaginary parts to cancel out. We then applied an inverse Fourier transform to generate the spatial noise pattern. The pattern was scaled in amplitude so that the root mean square contrast was 0.3. Values outside the interval [−0.98, 0.98] were clipped to the endpoints of this interval; approximately one pixel in 1,000 had to be clipped in this way. For correlated adaptation, each eye received the same noise image; for anticorrelated adaptation, the contrast was reversed between the eyes by multiplying each pixel in one eye's image by −1. Each eye's pattern was windowed with a sharp, circular envelope (diameter = 350 pixels) and surrounded by a black ring 4 pixels wide. Both the circular envelope and the black ring were antialiased to produce smooth edges. A small black fixation cross on an opaque white disk of diameter 8 pixels was inserted into the center of each adaptation image. The contrast values in the image were then converted to a luminance signal according to Equation 4 and gamma corrected for accurate presentation on the CRT monitor.

Apparatus

The setup was identical to that described in our previous work (May & Zhaoping, 2016). The left and right eyes' images were presented on the left and right sides of a 100-Hz Sony Trinitron CRT monitor screen, driven by a ViSaGe stimulus generator (Cambridge Research Systems, Rochester, UK), which produced images with a grayscale resolution of 14 bits/pixel. The background luminance was 54 cd/m^2, and the display was viewed through a mirror stereoscope that allowed each eye to view half the screen. Each screen pixel subtended 2.73 arcmin of visual angle.

Procedure

Before each block of 40 trials, the participant viewed a sequence of either 120 binocularly correlated or 120 anticorrelated noise patterns (0.5 s each). Then a black rectangle (identical to the black frame surrounding each test image—see later) appeared for 1 s to indicate that the first trial was about to start. Each trial began with a sequence of 10 binocular random-noise patterns of the same type as the initial 120 noise patterns (0.5 s each). The participant was then presented with a pair of binocular test images defined by either Equations 2 and 3 or Equations 7 and 8. The test images were presented for 0.4 s surrounded by a rectangular black border 4 pixels wide. After 0.4 s, the test images were overwritten with the uniform background gray, but the black border stayed on until the participant had responded. After the response, the screen was cleared to the

uniform background gray, and there was a short pause before the next trial began.

The participant pressed a left key to indicate Ant, Brad, male, happy, or rotated left; a right key indicated Dec, Matt, female, sad, or rotated right. For each category pair, half the trials had the left-key response corresponding to the summation channel's image (Equations 2 and 3) and the other half had the left-key response corresponding to the differencing channel's image (Equations 7 and 8); within each half of trials, half had the left eye viewing image I_1 and half had the right eye viewing image I_1. The temporal order of trials was random.

There were two "left" response keys and two "right" response keys, arranged in a single horizontal row. Participants were instructed to press the outer left or right key when they felt "reasonably confident" of their response and to press the inner left or right key otherwise. Figure 3 shows the data from all trials, regardless of confidence level. Supplementary Figure S1 shows the data for just the confident trials, and Supplementary Figure S2 shows the data for just the nonconfident trials. The separated data in the supplementary figures show the same pattern as the combined data in Figure 3, except that the aftereffect is larger for the confident responses, suggesting that when the adaptation was effective, it created a clear perception of the predicted face category.

For each category judgment, the summation channel's contrast α was adjusted before the experiment to make the binocular stimulus as low contrast as possible while still being clearly visible. The differencing channel's contrast β was almost always higher than α, to compensate for a previously reported bias to perceive the pattern presented to the summation channel (May & Zhaoping, 2016; May et al., 2012; Shadlen & Carney, 1986; Zhaoping, 2017). Individual differences between participants forced us to use a different β for each participant (all contrast values are given in Figure 3). To find β for each participant and category judgment, we ran a few pilot blocks with different β values. After finding a suitable β value for a particular category judgment, we usually ran 20 blocks for a participant for that category judgment, with the blocks alternating between correlated and anticorrelated adaptation (10 blocks of each, giving 400 trials for each data point in Figure 3). For one participant (diamonds in Figure 3), the Brad/Matt and left/right judgments were abandoned after 10 blocks (five of each adaptation type, giving 200 trials for each data point in Figure 3), as it was clear that this participant showed a negligible effect of adaptation on these judgments. For a further participant (circles in Figure 3), the male/female judgment consisted of 12 blocks (six of each adaptation type, giving 240 trials for each data point in Figure 3); this reduced number of trials for this

condition was necessitated because the participant did not have enough time to continue. For most face-category judgments, we ran all blocks for a participant for that judgment before switching to a different judgment. However, one participant (squares in Figure 3) interleaved blocks of the happy/sad and left/right judgment. Another participant (upward-pointing triangles) interleaved a replication of happy/sad (plotted in gray) with the left/right judgment.

The research was conducted in accordance with the Declaration of Helsinki. Written, informed consent was obtained from all observers, and approval of the study was obtained from the UCL Research Ethics Committee (project ID No. 6582/001).

Results

Figure 3 shows the results from the happy/sad judgment and from analogous judgments of identity (Brad Pitt/Matt Damon), gender (male/female), and head direction (left/right), using composite images constructed using the same principle as illustrated in Figure 1. Two participants also participated in a further identity judgment, this time with the British TV double act Ant and Dec. For each type of judgment, the effect of adaptation was highly significant: After correlated adaptation (which selectively desensitized the summation channel), participants mostly perceived the differencing channel's image; after anticorrelated adaptation (which selectively desensitized the differencing channel), participants mostly perceived the summation channel's image.

Discussion and conclusions

In this study, we extended our previous binocular adaptation paradigm to faces. We presented binocular test stimuli that delivered different face images to the binocular summation and differencing channels, and found that selective adaptation of one binocular channel would bias perception toward the other channel's face image.

As argued in the Introduction, the site of adaptation cannot possibly be the high-level cortical mechanisms that process faces. These adaptation effects must therefore result from inheritance of adaptation from earlier stages of processing. This shows that the inheritance of adaptation demonstrated physiologically in early stages of visual processing continues up to some of the highest levels of visual processing, where faces are analyzed. As noted earlier, in principle it was possible that these high-level mechanisms might have

compensated for low-level adaptation, to avoid corruption of the sensory signal (Xu et al., 2008). However, the face-adaptation effects that we report in this article must have resulted from inheritance of adaptation that was generated at much earlier stages of processing.

We have previously shown that selective adaptation of the binocular summation and differencing channels could influence judgments of simple attributes such as orientation (May & Zhaoping, 2016) and motion direction (May et al., 2012). However, these judgments of simple attributes could be mediated by mechanisms in the early stages of visual processing. Our new results provide conclusive evidence that adaptation at low-level sites can propagate up to the higher cortical areas where face-perception judgments are made, giving rise to face-adaptation effects, as argued by Dickinson and colleagues.

Keywords: vision, perception, psychophysics, face perception, adaptation, binocular integration

References

Afraz, A., & Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *Journal of Vision*, 9(10):10, 1–17, https://doi.org/10.1167/9.10.10. [PubMed] [Article]

Burger, W., & Burge, M. J. (2009). *Principles of digital image processing: Core Algorithms*. London, UK: Springer-Verlag.

Dhruv, N. T., & Carandini, M. (2014). Cascaded effects of spatial adaptation in the early visual system. *Neuron*, 81, 529–535, https://doi.org/10.1016/j.neuron.2013.11.025.

Dhruv, N. T., Tailby, C., Sokol, S. H., & Lennie, P. (2011). Multiple adaptable mechanisms early in the primate visual pathway. *The Journal of Neuroscience*, 31, 15016–15025, https://doi.org/10.1523/jneurosci.0890-11.2011.

Dickinson, J. E., Almeida, R. A., Bell, J., & Badcock, D. R. (2010). Global shape aftereffects have a local substrate: A tilt aftereffect field. *Journal of Vision*, 10(13):5, 1–12, https://doi.org/10.1167/10.13.5. [PubMed] [Article]

Dickinson, J., & Badcock, D. (2013). On the hierarchical inheritance of aftereffects in the visual system. *Frontiers in Psychology*, 4, 472, https://doi.org/10.3389/fpsyg.2013.00472.

Dickinson, J. E., Mighall, H. K., Almeida, R. A., Bell, J., & Badcock, D. R. (2012). Rapidly acquired shape and face aftereffects are retinotopic and local in origin. *Vision Research*, 65, 1–11, https://doi.org/10.1016/j.visres.2012.05.012.

Galton, F. (1878). Composite portraits. *Nature*, 18, 97–100, https://doi.org/10.1038/018097a0.

Hills, P. J., & Lewis, M. B. (2012). FIAEs in famous faces are mediated by type of processing. *Frontiers in Psychology*, 3, 256, https://doi.org/10.3389/fpsyg.2012.00256.

Hole, G. (2011). Identity-specific face adaptation effects: Evidence for abstractive face representations. *Cognition*, 119, 216–228, https://doi.org/10.1016/j.cognition.2011.01.011.

Kohn, A., & Movshon, J. A. (2003). Neuronal adaptation to visual motion in area MT of the macaque. *Neuron*, 39, 681–691, https://doi.org/10.1016/S0896-6273(03)00438-0.

Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience*, 7, 764–772, https://doi.org/10.1038/nn1267.

Larsson, J., & Harrison, S. J. (2015). Spatial specificity and inheritance of adaptation in human visual cortex. *Journal of Neurophysiology*, 114, 1211–1226, https://doi.org/10.1152/jn.00167.2015.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, 4, 89–94, https://doi.org/10.1038/82947.

Li, Z., & Atick, J. J. (1994). Efficient stereo coding in the multiscale representation. *Network: Computation in Neural Systems*, 5, 157–174.

May, K. A., & Zhaoping, L. (2016). Efficient coding theory predicts a tilt aftereffect from viewing untilted patterns. *Current Biology*, 26, 1571–1576, https://doi.org/10.1016/j.cub.2016.04.037.

May, K. A., Zhaoping, L., & Hibbard, P. B. (2012). Perceived direction of motion determined by adaptation to static binocular images. *Current*

Biology, *22*, 28–32, https://doi.org/10.1016/j.cub.2011.11.025.

Morgan, M., Dillenburger, B., Raphael, S., & Solomon, J. A. (2012). Observers can voluntarily shift their psychometric functions without losing sensitivity. *Attention, Perception, & Psychophysics*, *74*(1), 185–193, https://doi.org/10.3758/s13414-011-0222-7.

Nicholson, B. J. (1985). On the *F*-distribution for calculating Bayes credible intervals for fraction nonconforming. *IEEE Transactions on Reliability*, *R-34*, 227–228.

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., … Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, *394*, 884–887, https://doi.org/10.1038/29772.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342, https://doi.org/10.1007/bf00239352.

Priebe, N. J., Churchland, M. M., & Lisberger, S. G. (2002). Constraints on the source of short-term motion adaptation in macaque area MT. I. The role of input and intrinsic mechanisms. *Journal of Neurophysiology*, *88*, 354–369, https://doi.org/10.1152/jn.00852.2001.

Rolls, E. T., & Baylis, G. C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research*, *65*, 38–48, https://doi.org/10.1007/bf00243828.

Shadlen, M., & Carney, T. (1986, Month DD). Mechanisms of human motion perception revealed by a new cyclopean illusion. *Science*, *232*, 95–97.

Solomon, S. G., Peirce, J. W., Dhruv, N. T., & Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron*, *42*, 155–162, https://doi.org/10.1016/S0896-6273(04)00178-3.

Troy, J. B., & Enroth-Cugell, C. (1993). X and Y ganglion cells inform the cat's brain about contrast in the retinal image. *Experimental Brain Research*, *93*, 383–390, https://doi.org/10.1007/bf00229354.

Vakli, P., Németh, K., Zimmer, M., Schweinberger, S. R., & Kovács, G. (2012). Face distortion aftereffects evoked by featureless first-order stimulus configurations. *Frontiers in Psychology*, *3*, 566, https://doi.org/10.3389/fpsyg.2012.00566.

Watson, T. L., & Clifford, C. W. G. (2003). Pulling faces: An investigation of the face-distortion aftereffect. *Perception*, *32*, 1109–1116, https://doi.org/10.1068/p5082.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*, 557–561, https://doi.org/10.1038/nature02420.

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1702–1725, https://doi.org/10.1098/rstb.2010.0360.

Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review*, *6*, 647–653, https://doi.org/10.3758/BF03212974.

Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the cortical hierarchy: Low-level curve adaptation affects high-level facial-expression judgments. *The Journal of Neuroscience*, *28*, 3374–3383, https://doi.org/10.1523/jneurosci.0182-08.2008.

Xu, H., Liu, P., Dayan, P., & Qian, N. (2012). Multi-level visual adaptation: Dissociating curvature and facial-expression aftereffects produced by the same adapting stimuli. *Vision Research*, *72*, 42–53, https://doi.org/10.1016/j.visres.2012.09.003.

Yamashita, J. A., Hardy, J. L., De Valois, K. K., & Webster, M. A. (2005). Stimulus selectivity of figural aftereffects for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 420–437, https://doi.org/10.1037/0096-1523.31.3.420.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research*, *41*, 2979–2994, https://doi.org/10.1016/S0042-6989(01)00202-4.

Zhaoping, L. (2017). Feedback from higher to lower visual areas for visual recognition may be weaker in the periphery: Glimpses from the perception of brief dichoptic stimuli. *Vision Research*, *136*, 32–49, https://doi.org/10.1016/j.visres.2017.05.002.