

The Role of Background Statistics in Face Adaptation

Jianhua Wu,^{1*} Hong Xu,^{1*} Peter Dayan,² and Ning Qian¹

¹Departments of Neuroscience and Physiology & Cellular Biophysics, Columbia University, New York, New York 10032, and ²Gatsby Computational Neuroscience Unit, University College London, London WC1N 3AR, United Kingdom

Cross-adaptation is widely used to probe whether different stimuli share common neural mechanisms. For example, that adaptation to second-order stimuli usually produces little aftereffect on first-order stimuli has been interpreted as reflecting their separate processing. However, such results appear to contradict the cue-invariant responses of many visual cells. We tested the novel hypothesis that the null aftereffect arises from the large difference in the backgrounds of first- and second-order stimuli. We created second-order faces with happy and sad facial expressions specified solely by local directions of moving random dots on a static-dot background, without any luminance-defined form cues. As expected, adaptation to such a second-order face did not produce a facial-expression aftereffect on the first-order faces. However, consistent with our hypothesis, simply adding static random dots to the first-order faces to render their backgrounds more similar to that of the adapting motion face led to a significant aftereffect. This background similarity effect also occurred between different types of first-order stimuli: real-face adaptation transferred to cartoon faces only when noise with correlation statistics of real faces or natural images was added to the cartoon faces. These findings suggest the following: (1) statistical similarities between the featureless backgrounds of the adapting and test stimuli can influence aftereffects, as in contingent adaptation; (2) weak or null cross-adaptation aftereffects should be interpreted with caution; and (3) luminance- and motion-direction-defined forms, and local features and global statistics, converge in the representation of faces.

Introduction

A central question in sensory processing is whether different classes of stimuli for the same sensory attribute share a common neural representation. The standard signature of such sharing is the transfer of an aftereffect in cross-adaptation. For vision, this question has been prominently raised with respect to the comparison between first-order, luminance-defined stimuli and second-order stimuli defined by variations in contrast, motion, texture, phase, and spatial or temporal frequency, etc., without modulation of mean luminance. The common psychophysical and neuroimaging finding that adaptation to second-order stimuli does not substantially transfer to first-order stimuli (Nishida et al., 1997; Larsson et al., 2006; Ashida et al., 2007; Schofield et al., 2007) has duly been interpreted as indicating that the two classes of stimuli are processed separately.

Such separation, however, appears to be at odds with cue-invariant responses found at many stages along the visual hierarchy. For example, some V1 cells and a larger fraction of V2 cells have similar tuning to first-order and second-order (illusory-contour) orientations (von der Heydt et al., 1984; Sheth et al.,

1996). Likewise, subsets of MT and MSTd cells show similar tuning to first- and second-order motion patterns (Albright, 1992; Geesaman and Andersen, 1996). Moreover, shape selectivity does not depend on whether the shapes are defined by first- or second-order cues (Sary et al., 1993; Grill-Spector et al., 1998). If such cells contribute significantly to adaptation, then aftereffects should transfer, at least partially, from second- to first-order stimuli. Conversely, if they do not contribute, then adaptation seems an inadequate probe for shared neural mechanisms.

One hitherto unexplored confounding factor is that, with the notable exception of the one study (Georgeson and Schofield, 2002) that did find significant interorder cross-adaptation transfer, first- and second-order stimuli, by their construction, often have different background patterns. For instance, a typical first-order orientation stimulus is a black bar on a uniform white background; however, its second-order counterpart might involve flickering dots in a virtual rectangle embedded in a larger background of static dots. Featureless backgrounds could be important since, for instance, their size makes them contribute substantially to the statistics of images, and visual adaptation is highly sensitive to such statistics (Field, 1987; Atick and Redlich, 1990; Barlow, 1990; Dan et al., 1996; Wainwright, 1999; Brenner et al., 2000; Simoncelli and Olshausen, 2001; Zhaoping, 2006; Schwartz et al., 2007). However, theories considering image statistics focus on predicting visual receptive fields rather than cross-adaptation aftereffects, whereas theories treating aftereffects consider only statistics of features but not backgrounds.

Therefore, we investigated the role of background statistics in the transfer of the facial-expression aftereffect between different classes of faces. We chose face stimuli because cue convergence may be more pronounced at higher levels of processing; we will

Received May 19, 2009; revised July 15, 2009; accepted Aug. 12, 2009.

This work was supported by National Institutes of Health Grant EY016270 (N.Q.) and by the Gatsby Charitable Foundation and the Biotechnology and Biological Sciences Research Council, Engineering and Physical Sciences Research Council, and Wellcome Trust Grant BB/E002536/1 (P.D.). We are very grateful to an anonymous reviewer for suggesting that we relate our results to contingent aftereffects and to Richard Lipkin for editorial assistance.

*J.W. and H.X. contributed equally to this work.

Correspondence should be addressed to Dr. Ning Qian, Department of Neuroscience, Columbia University/New York State Psychiatric Institute, Kolb Annex, Room 519, 1051 Riverside Drive, Box 87, New York, NY 10032. E-mail: nq6@columbia.edu.

H. Xu's present address: Division of Psychology, Nanyang Technological University, 637332, Singapore.

DOI:10.1523/JNEUROSCI.2346-09.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/2912035-10\$15.00/0

report similar studies for lower-level stimuli separately. We found that background statistics can be critically important for interactions not only between first- and second-order stimuli but also between different types of first-order stimuli. These findings have implications for the function and mechanisms of adaptation and the representation of faces.

Materials and Methods

Subjects

A total of eight subjects consented to participate in the experiments of this study. Among them, two were experimenters, and the rest were naive to the purpose of the study. All subjects had normal or corrected to normal vision. Each experimental condition had four subjects, two of whom were naive. All subjects showed highly consistent results. The study was approved by the Institutional Review Board of the New York State Psychiatric Institute.

Apparatus

The visual stimuli were presented on a 21 inch ViewSonic P225f monitor controlled by a Macintosh G4 computer. The vertical refresh rate was 100 Hz, and the spatial resolution was 1024×768 pixels. In a dimly lit room, subjects viewed the monitor from a distance of 75 cm, using a chin rest to stabilize head position. Each pixel subtended 0.029° at this distance. All luminance values cited in the paper were measured with a Minolta LS-110 photometer. All experiments were run in Matlab with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Visual stimuli

A black (0.23 cd/m^2) fixation cross was always shown at the center of the white (56.2 cd/m^2) screen. It consisted of two orthogonal line segments, each 0.29° in length and 0.06° in width. All stimuli were grayscale. They included second-order cartoon faces with motion gradient-defined facial expressions, first-order cartoon faces with luminance-defined facial expressions, and first-order real faces derived from the Ekman Pictures of Facial Affect (PoFA) database (Ekman and Friesen, 1976). Synthetic noises were used as backgrounds for some of the face stimuli (see below). Each stimulus was shown with the fixation cross at its center.

Motion gradient cartoon faces. We generated these dynamic, second-order faces in a $3.7^\circ \times 3.7^\circ$ (129×129 pixels) square containing black random dots on a white background. Each dot had a size of 1 pixel (0.029°). The dot density was 15%. On this random-dot field, we defined a virtual ring for the face outline, two virtual disks for the eyes, and a horizontally oriented virtual rectangle for the mouth (Fig. 1a). The outline ring had inner and outer radii of 1.77° and 1.48° , respectively, containing 500 dots on average. Each eye disk had a radius of 0.17° , containing 17 dots on average. The mouth rectangle was $0.29^\circ \times 1.28^\circ$, containing 66 dots on average. The mouth center was 0.64° below the face center. The distance between the eye and mouth levels was 1.22° . The center-to-center distance between the eyes was 0.85° . The dots within the face-outline ring and the eye disks randomly flickered at 10 Hz. This was achieved by letting each dot have a lifetime of 100 ms (10 frames) and get replotted to a random location within its original region when its lifetime was reached. The initial lifetimes of different dots were randomized to avoid synchronous flicker. The dots within the mouth rectangle had different motion gradient patterns to produce sad and happy expressions, as explained in the text (Fig. 1b,c). The motion along each vertical line of the mouth rectangle was created by moving a window of 1 pixel width on a large field of random dots at the desired velocity; this avoided dots' jumping between opposite edges of the rectangle. To generate a set

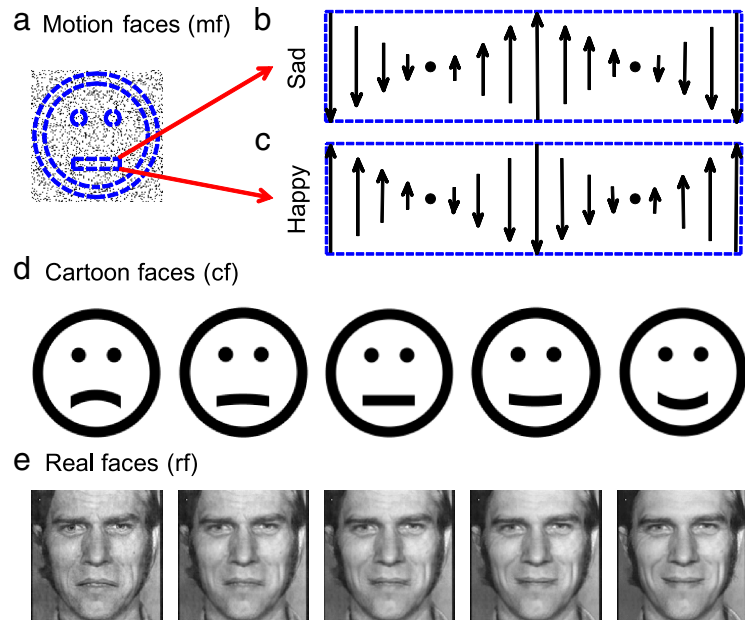


Figure 1. Examples of face stimuli used in this study. *a*, An illustration for second-order motion faces with motion gradient-defined expressions. Dashed blue curves and lines define virtual regions for the face outline, eyes, and mouth. Dots inside the face outline and eye regions flickered at 10 Hz. *b*, *c*, Dots inside the mouth rectangle carried opposite motion gradient patterns to produce sad and happy expressions. The vertical arrows indicate the velocity of the dots in that column. *d*, Static cartoon faces generated by our anti-aliasing program. The mouth curvature varied from concave to convex to produce a spectrum of sad to happy expressions. *e*, Real faces derived from two images from the Ekman PoFA database with MorphMan 4.0.

of motion faces with expressions varying gradually from sad to happy, we let a fraction of the dots (noise dots) in the mouth rectangle flicker at 10 Hz. These stimuli are referred to as motion faces.

Static cartoon faces. As in a previous study (Xu et al., 2008), we generated black cartoon faces on white background with an anti-aliasing method. The dimensions of the face outline ring and eye disks exactly matched those of the motion faces (Fig. 1d). The length and thickness of the mouth curve also matched those of the motion faces. However, unlike the motion faces whose mouths were a flat rectangle, the mouths of the static cartoon faces had curvature varying from concave to convex ($-0.34, -0.25, -0.17, -0.08, 0, 0.08, \text{ and } 0.17$ in the units of $1/\sigma$) to produce a spectrum of sad to happy expressions. The positions of the eyes and mouth within the static cartoon faces were also identical to those in the motion faces.

Real-face images. These stimuli were derived from images from Ekman PoFA database (Ekman and Friesen, 1976) in the same way as a previous study (Xu et al., 2008), except that we cropped and scaled the image size to $3.7^\circ \times 3.7^\circ$ and presented them at a higher contrast (mean mouth contrast, 0.27). We applied MorphMan 4.0 (STOIK Imaging) to a sad and a happy face of the same person to generate 21 images with the proportion of the happy face varying from 0 (saddest) to 1 (happiest) in steps of 0.05 (Fig. 1e). Those with proportions equal to 0, 0.15, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, and 0.7 were used in the experiments.

Noise backgrounds with various correlation statistics. We generated $3.7^\circ \times 3.7^\circ$ noise patterns as backgrounds for static cartoon faces in the last experiment. We first computed correlation statistics of real faces by calculating the average amplitude spectrum of the front-view faces in the Ekman PoFA database (Ekman and Friesen, 1976) and Karolinska Directed Emotional Faces (KDEF) database (Lundqvist et al., 1998). We used the standard procedure of first computing the 2D Fourier transform of the images and then combining the amplitude spectrum across different orientations to obtain a 1D spectrum. We found that the spectrum follows the $1/f^k$ law with $k = 1.63$, where f is the spatial frequency.

We then generated $1/f^k$ noise patterns with k equal to 0, 0.5, 1.0, 1.5, 2.0, and 2.5. The $k = 0$ noise had a flat spectrum and we generated it as spatial white noise. For other k values, we inverse transformed the $1/f^k$ amplitude spectrum and a random phase spectrum to the spatial do-

main. Finally, we applied the standard histogram equalization procedure (MATLAB `histeq` function) to equalize the first-order distributions of the noises to that of the saddest real face. Therefore, noises with different k values all had the same first-order distributions but different second-order correlation structures.

Procedures

We used the method of constant stimuli and the two-alternative forced-choice paradigm in all experiments. Subjects received no feedback on their performances at any time. We measured facial expression perception in the motion faces (mf) and in the controls whose mouth region was frame randomized to destroy motion (mf_r). These conditions did not involve adaptation and are denoted as 0-mf and 0-mf_r. For the 0-mf condition, we randomly interleaved the upside-down happiest and saddest motion faces as catch trials for the two naive subjects to ensure that they judged facial expression, instead of motion direction (see Results). We also cross-adapted the motion faces and the static cartoon faces (cf). The adapting stimulus was either the saddest mf or the saddest cf and the test stimuli were either the set of the motion faces or the set of the static cartoon faces. The four possible combinations are denoted as mf-mf, mf-cf, cf-mf, and cf-cf, where, for example, mf-cf represents the condition with the saddest mf as the adapting stimulus and the cf set as the test stimuli. Similarly, we measured cross-adaptation aftereffects between the motion faces and the real faces (rf), resulting in four possible conditions: mf-mf, mf-rf, rf-mf, and rf-rf. We also included baseline conditions without adaptation for cf and rf, and they are denoted as 0-cf and 0-rf. (The baseline condition for mf was the same as the 0-mf condition mentioned above.) In total, there were 11 different conditions (0-mf, 0-mf_r, 0-cf, 0-rf, mf-mf, cf-cf, rf-rf, mf-cf, cf-mf, mf-rf, and rf-mf).

The 11 conditions were run in separate blocks with two blocks per condition. Although in Results we described these conditions in a specific order, they were randomized for each subject. Over the two blocks for each condition, each test stimulus was repeated 20 times. The trials for different test stimuli in a block were also randomized. There was a break of at least 10 min after each adaptation block to avoid carryover of the aftereffects to the next block. Data collection for each block started after subjects had sufficient practice trials (~10–20) to feel comfortable with the task.

Subjects started each block of trials by fixating the central cross and pressing the space bar. After 500 ms, for each adaptation block the adapting stimulus appeared for 4 s. After a 500 ms interstimulus interval (ISI), a test face appeared for 100 ms if it was a static cartoon face or real face and for 1 s if it was a motion face. The longer duration for the motion face was needed for subjects to perform the task well. For the baseline blocks without adaptation, only a test stimulus was shown in each trial. A 50 ms beep was then played to remind subjects to report their perception of the test stimulus. Subjects had to press the “A” or “S” key to report happy or sad expression, respectively. After a 1 s intertrial interval, the next trial began.

After the completion of the above 11 conditions, we ran additional conditions to study how adaptation to the contrast-matched cartoon face (cf_c) and modified real face (rf_m) affected the perception of the motion faces (the cf_c-mf and rf_m-mf conditions). cf_c was the saddest static cartoon face, with its foreground and background luminance values set to the means of the mouth and surrounding areas of the saddest real face. rf_m was the saddest real face, with the mouth curve of cf_c pasted over. Moreover, we investigated whether motion-face adaptation could produce an aftereffect on the static cartoon faces with random dots added (cf_d) to make their background more similar to the motion face (the mf-cf_d and 0-cf_d conditions). We also ran two control experiments. The first used test cartoon faces whose uniform background luminance matched that of the cf_d stimuli but without any dots (the mf-cf_i and 0-cf_i conditions). The second experiment repeated the mf-cf_d condition but subjects were asked to judge motion direction (up or down) around the midpoint of the mouth curves of the test static faces (the mf-cf_d-dir condition). The rationale for running these seven conditions was explained in Results.

Finally, we cross-adapted real faces and static cartoon faces. We started with four adaptation conditions, namely, cf-cf, cf-rf, rf-cf, and rf-rf

(where rf-cf, for example, means adapting to the saddest real face and testing on the cartoon faces), and two baseline conditions, 0-cf and 0-rf. We then repeated the rf-cf and 0-cf conditions after adding $1/f^k$ noises to the background of the test cartoon faces (the rf-cf_k and 0-cf_k conditions). We used $k = 0, 0.5, 1.0, 1.5, 2.0,$ and 2.5 in different blocks with randomized order. The noise in each trial was generated online. Therefore, although noise backgrounds in the same block all had the same correlation statistics (same k), there was no repetition of a specific sample. We also included a uniform background whose luminance equaled the mean of the other backgrounds. This corresponds to a very large k and is labeled as $k = \infty$. The rationale for running these 20 conditions was explained in Results.

Data analysis

For each condition, the data were sorted into fraction of “happy” responses to each test stimulus. The test stimuli were parameterized according to the fraction of the signal dots in the motion faces, the mouth curvature of the static cartoon faces, or the proportion of the happy face in the morphed real faces. The fraction of “happy” responses was then plotted against the test stimulus, and the resulting psychometric curve was fitted with a sigmoidal function of the form $f(x) = 1/[1 + e^{-a(x-b)}]$, where a determines the slope and b gives the test stimulus parameter corresponding to the 50% point of the psychometric function [the point of subjective equality (PSE)]. An aftereffect is measured by the difference between the PSEs of the adaptation condition and the corresponding baseline condition. To determine whether an aftereffect was significant, we calculated the p value by comparing subjects’ PSEs of the adaptation condition against those of the corresponding baseline condition via a two-tailed paired t test. For the 0-mf_r and mf-cf_d-dir conditions, we determined whether the slopes of the psychometric curves were significantly different from zero via a two-tailed t test. In the last experiment, in which the k value of the noise background was varied, we used ANOVA to test the dependence on k .

Results

We first describe a novel class of second-order faces in which facial expressions are either sad or happy depending on local directions of motion. We then present our cross-adaptation studies using these second-order faces and conventional first-order cartoon and real faces and demonstrate the importance of statistical similarity between the featureless backgrounds of the adapting and test stimuli.

Facial expressions solely specified by motion gradients in second-order faces

We programmed second-order cartoon faces with motion gradient-defined expressions (Fig. 1*a*). On a white screen, we generated black random dots with a 15% dot density over a $3.7^\circ \times 3.7^\circ$ square. We then defined a virtual ring for the face outline, two virtual disks for the eyes, and a horizontally oriented virtual rectangle for the mouth. Unlike first-order stimuli, these regions were not drawn with a luminance different from the background. Instead, they were virtual regions within which dots were dynamic and thus perceptually segregated from the background static dots. The dots within the outline ring and the eye disks randomly flickered at 10 Hz. The dots within the mouth rectangle had different motion gradient patterns to produce sad and happy expressions. For the sad expression (Fig. 1*b*), the dots along the middle vertical line of the rectangle moved upward and those at the two ends of the rectangle moved downward at a speed of $2.9^\circ/s$. The velocity of other dots in the rectangle was determined by a linear interpolation. Owing to the motion mislocalization effect (Ramachandran and Anstis, 1990; De Valois and De Valois, 1991; Fu et al., 2004), this motion gradient pattern made the mouth rectangle look concave even though it was physically flat. For the happy expression, the direction of each dot in the mouth rectangle was reversed (Fig. 1*c*). The first supplemental file, avail-

able at www.jneurosci.org as supplemental material, contains two movies showing that indeed these stimuli appear sad and happy, respectively. Playing the movies backward to reverse the directions of motion exchanges the sad and happy expressions. Stopping the movies at any frame causes both the faces and their expressions to vanish. When we randomized the frames of the mouth rectangle (see the second supplemental file, available at www.jneurosci.org as supplemental material), the perceived differences between sad and happy faces disappeared, further proving the lack of static form cues for the expressions.

To conduct formal psychophysical experiments, we generated a set of faces with expressions that gradually changed from sad to happy by letting a fraction of the dots (signal dots) within the mouth rectangle follow the above-prescribed motion gradients while the remaining dots (noise dots) randomly flickered at 10 Hz. We use negative and positive fractions to label the motion gradient patterns for the sad and happy expressions, respectively. Thus, the original sad and happy faces with 100% signal dots in the mouth region have signal dot fractions of -1 and 1 , respectively. A fraction of 0 means that all dots are noise dots. We generated a total of nine stimuli with fractions from -1 to 1 , in steps of 0.25 . Subjects initiated a block of trials after fixating on a central cross. In each trial, a stimulus was pseudorandomly selected and shown centered at fixation. Subjects reported whether the perceived expression was sad or happy via a key press. At no point was feedback given in this or the other experiments.

The psychometric data for four subjects (two of them naive) are shown as the blue curves (the 0-mf condition) in Figure 2*a–d*. As the fraction of signal dots varied from -1 to 1 , the fraction of “happy” responses increased gradually from 0 to 1 for all subjects, indicating that the perceived expressions varied from sad to happy in an orderly manner. One might raise the possibility that subjects did not really perceive any facial expressions but simply judged local motion directions. This is unlikely for three reasons. First, none of the subjects performed a motion direction task under this condition so there was little chance of confusion. Second, if the subjects could not see facial expression and decided to use motion direction as a substitute, it would be hard to explain why all subjects happen to choose the same motion pattern to indicate happy or sad expressions. Third, when they were given the instructions, all subjects understood the task immediately and found it to be straightforward. To rule out this possibility directly, for the two naive subjects, we randomly interleaved catch trials in which the stimuli were the saddest and happiest faces (signal dot fractions of -1 and 1) shown upside-down. When the face with a signal dot fraction of 1 (or -1) was inverted, its motion gradient pattern was identical to that of the upright face with a signal dot fraction of -1 (or 1). If the subjects judged local motion directions, they would consider the inverted saddest face as “happy” and the inverted happiest face as “sad.” The data, marked as inverted happy or sad face in Figure 2, *a* and *b*, show the opposite results, indicating that the subjects indeed judged

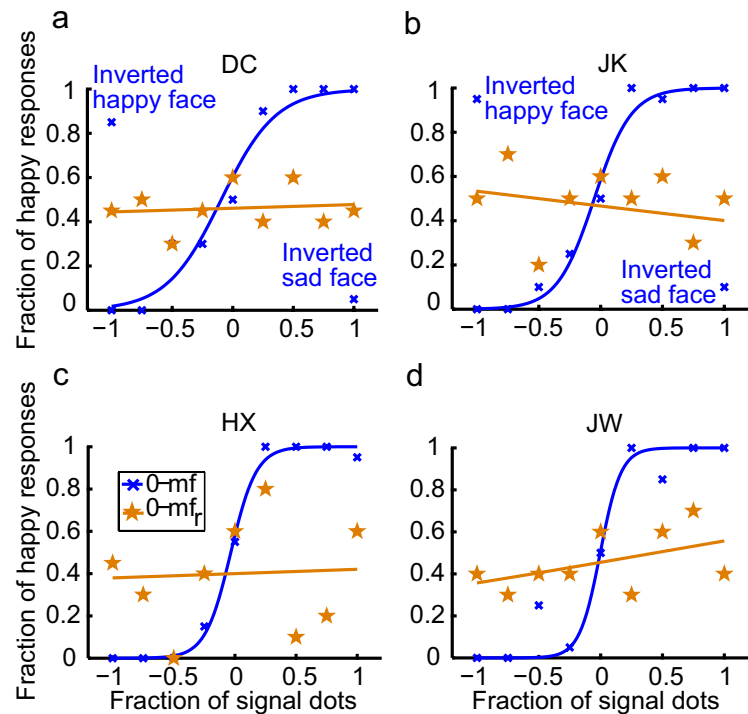


Figure 2. Facial-expression perception with the motion faces. *a–d*, Psychometric functions from two naive subjects (DC, JK) and two experimenters (HX, JW). For each subject, the perceived expression varied gradually from sad to happy as the fraction of the signal dots varied from -1 to 1 (blue curve, the 0-mf condition). The perception of expression was destroyed with randomized motion frames (brown curve, the 0-mf_r condition). For the two naive subjects, catch trials with the inverted saddest and happiest motion faces were shown to indicate that they judged facial expressions instead of motion directions (see Results for details).

the faces’ expressions. The fact that the catch trials did not produce chance performances (0.5) also suggests that for simple cartoon faces, inversion does not destroy the perception of facial expressions (Xu et al., 2008).

To ensure that the perceived facial expressions in the above experiment were completely specified by motion instead of by static form cues, we conducted a control experiment that randomized the motion frames of the mouth rectangle for all nine face stimuli. The results of the subjects’ judgments are shown as the brown curves (the 0-mf_r condition) in Figure 2*a–d*. The slopes of these psychometric curves do not significantly differ from zero (mean, 0.023 ; $p = 0.47$), indicating that the subjects could not distinguish between sad and happy expressions once the coherent motion gradients were destroyed by randomization. We conclude that there was no first-order form cue in the original stimuli.

Cross-adaptation between the second-order motion faces and first-order static cartoon faces and real faces

After establishing facial-expression perception in the second-order motion faces, we cross-adapted them with first-order, luminance-defined faces to gain insight into interactions between different cues in face processing. Previous studies on cross-adaptation between first- and second-order stimuli have been limited to low-level stimuli (Nishida et al., 1997; Larsson et al., 2006; Ashida et al., 2007; Schofield et al., 2007).

As in our previous study (Xu et al., 2008), we generated first-order static cartoon faces and real faces with expressions varying from sad to happy; examples are shown in Figure 1, *d* and *e*. The physical dimensions of the static cartoon faces exactly matched those of the motion faces except that the sad to happy expressions were generated by luminance-defined curvature rather than mo-

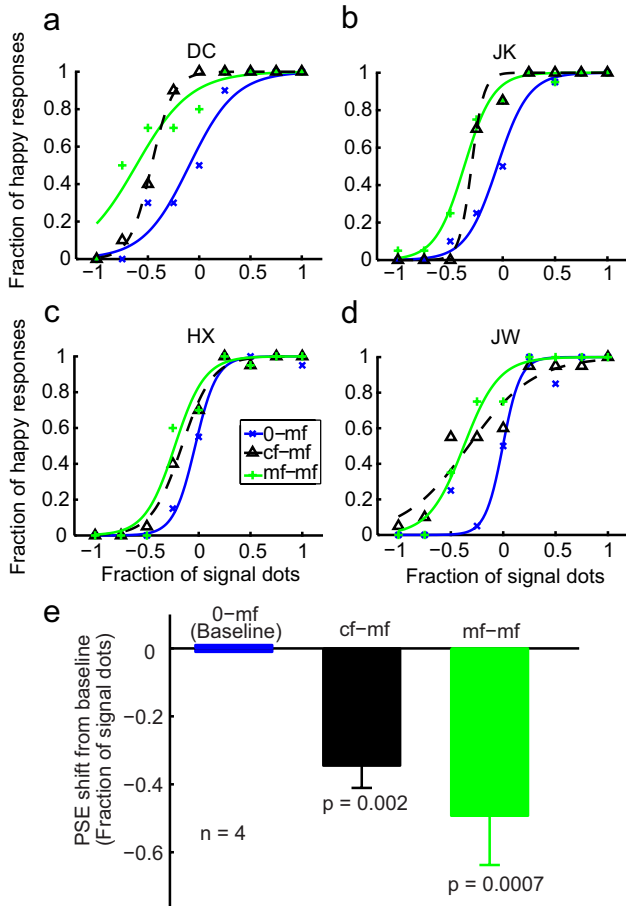


Figure 3. The effect of cartoon adaptation on the perceived expression of the motion faces. **a–d**, Psychometric functions from two naive subjects (DC, JK) and two experimenters (HX, JW). The test stimuli were always the mf. The adapting stimuli varied with conditions as follows: 0-mf, no-adaptation baseline (blue, copied from Fig. 2); cf-mf, adaptation to the saddest cartoon face (dashed black); mf-mf, adaptation to the saddest motion face (green). **e**, Summary of all four subjects' data. For each condition, the average PSE relative to the baseline condition was plotted, with the error bar representing \pm SEM. The *p* value shown for each condition in the figure was calculated against the baseline condition using a two-tailed paired *t* test.

tion gradients. We also closely matched the motion faces and the real faces, paying special attention to align the mouth positions of the real faces to that of the motion faces, since the mouth is particularly important for facial expressions (Gosselin and Schyns, 2001; Xu et al., 2008).

We first report how the second-order motion faces and the first-order static cartoon faces interacted. We used either the saddest mf or the saddest cf we generated as the adapting stimulus and then our set of mf or cf faces as the test stimuli. The total of four possible combinations are denoted as mf-mf, mf-cf, cf-cf, and cf-mf conditions, where, for example, cf-mf means adapting to the saddest cartoon face and testing on the motion faces. We also included baseline conditions without adaptation for the motion faces and static cartoon faces, denoting them as 0-mf (already described above) and 0-cf.

The psychometric data from four subjects (two of them naive) for the 0-mf, mf-mf, and cf-mf conditions are shown in Figure 3a–d as blue, green, and dashed black curves, respectively. Similarly, the data for the 0-cf, cf-cf, and mf-cf conditions are shown in Figure 4a–d as blue, green, and dashed black curves, respectively. As expected, the two within-category adaptation conditions (mf-mf and cf-cf) generated large facial-expression

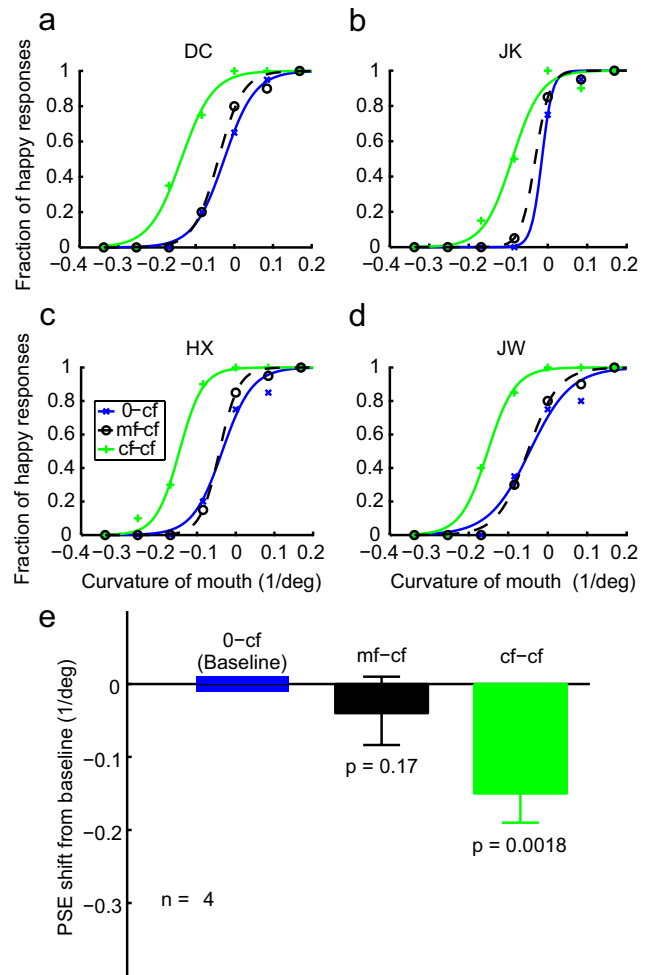


Figure 4. The effect of motion-face adaptation on the perceived expression of the static cartoon faces. **a–d**, Psychometric functions from two naive subjects (DC, JK) and two experimenters (HX, JW). The test stimuli were always the static cf. The adapting stimuli varied with conditions as follows: 0-cf, no-adaptation baseline (blue); cf-cf, adaptation to the saddest cartoon face (green); mf-cf, adaptation to the saddest motion face (dashed black). **e**, Summary of all four subjects' data.

aftereffects, indicated by the shifts of the psychometric curves from the corresponding baseline conditions (0-mf and 0-cf). The leftward shifts mean that subjects perceived happy expressions more often after adapting to sad faces, consistent with previous face adaptation experiments (Leopold et al., 2001; Webster et al., 2004; Xu et al., 2008). For the cross-adaptation conditions (cf-mf and mf-cf), the results were asymmetric: first-order cartoon-face adaptation generated a significant aftereffect on the second-order motion faces (cf-mf), yet second-order motion-face adaptation failed to generate an aftereffect on the first-order cartoon faces (mf-cf). Similar asymmetric interactions have been reported previously for low-level stimuli (Smith et al., 2001; Ellemberg et al., 2004; Schofield et al., 2007).

To quantify the aftereffects and summarize the results from all four subjects, we determined the PSE—the *x*-axis value corresponding to 50% “happy” responses—for each psychometric curve of each subject. Figure 3e shows the mean PSEs of the mf-mf (green bar) and cf-mf (black bar) conditions relative to the baseline 0-mf condition. Similarly, Figure 4e shows the mean PSEs of the cf-cf (green bar) and mf-cf (black bar) conditions relative to the baseline 0-cf condition. A negative value means a leftward shift of the psychometric curve from the baseline. The

error bars indicate ± 1 SEM. The asymmetry in the aftereffect is apparent, with the cf-mf, but not the mf-cf, condition being significantly different from the corresponding baseline (Figs. 3e and 4e, *p* values).

We also cross-adapted first-order real faces and the second-order motion faces and found little aftereffect transfer in either direction (see the third supplemental file, available at www.jneurosci.org as supplemental material). Similar results have also been reported for low-level stimuli (Nishida et al., 1997; Larsson et al., 2006; Ashida et al., 2007). Note, in particular, that the first-order cartoon face, but not the first-order real face, produced an aftereffect on the second-order motion faces. This is not attributable to different contrasts or mouth shapes of the stimuli (see the third supplemental file, available at www.jneurosci.org as supplemental material).

The role of background similarity in determining aftereffect transfer from second- to first-order faces

As mentioned in the Introduction, to understand the above results and the role of cue-invariant cells in cross-adaptation, we considered a new contingent factor that could contribute to the transfer of an aftereffect, namely, the similarity between the backgrounds of the adapting and test stimuli. By background we mean the luminance distributions of the largely featureless areas surrounding the facial features. We hypothesized that strong aftereffects are induced only when adapting and test stimuli have sufficiently similar backgrounds. The hypothesis implies the surprising prediction that although featureless backgrounds cannot produce a facial expression aftereffect by themselves, they will significantly influence the aftereffect. We designed experiments to test this prediction.

In Figure 4, we showed that motion-face adaptation failed to produce an aftereffect on the test static cartoon faces (the mf-cf condition). By the background-similarity hypothesis, we expected to observe the aftereffect in the mf-cf condition simply by adding random dots to the test cartoon faces to make them more similar to the adapting motion face. Importantly, this experiment could also address the question of whether it is possible for a second-order face to produce an aftereffect on first-order faces. We therefore generated a new set of test cartoon faces (cf_d) (Fig. 5a) by adding random dots to the original set shown in Figure 1d. The dot density was the same as for the motion faces (15%). The dots were added online for each trial so that no dot pattern was repeated across trials. We ran the motion-face adaptation condition (mf- cf_d) and the no-adaptation baseline condition (0- cf_d) in exactly the same way that we ran the original mf-cf and 0-cf conditions. The data from four subjects (two of them naive) are shown as dashed and solid red curves in Figure 6a–d. Motion-face adaptation indeed produced a significant aftereffect on the cartoon faces with added dots (Fig. 6e, hatched red bar), supporting the notion that background similarity influences the aftereffect. The result suggests that the lack of aftereffect in the original mf-cf condition (Fig. 4) is not entirely because the adapting motion face

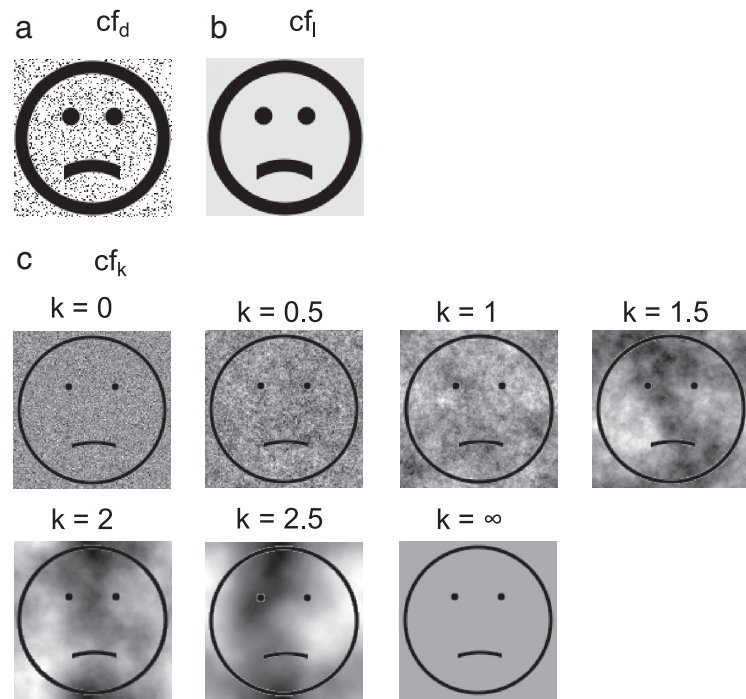


Figure 5. Manipulations of background similarity between adapting and test faces. Each face set contained expressions ranging from sad to happy, but only the saddest one is shown here. **a**, Cartoon face set with random dots added (cf_d). Random dots of 15% density were added to the original static cartoon face set in Figure 1d. **b**, Luminance-matched cartoon face set (cf_l). The background luminance was matched to the mean background luminance of cf_d set. **c**, Cartoon face sets with $1/f^k$ noise backgrounds, with $k = 0, 0.5, 1, 1.5, 2, 2.5$, respectively. Noises with different k values had the same first-order luminance distribution but different second-order correlation structures. A uniform background was also included and labeled as $k = \infty$.

is not as salient as the test cartoon faces. It further suggests that second-order stimuli can produce aftereffects on first-order stimuli when their backgrounds are more similar.

We excluded two alternative explanations. The first depends on the fact that the added black dots reduced the mean background luminance of the cf_d set (47.6 cd/m^2) compared with the original cartoon faces (56.2 cd/m^2), resulting in a slightly lower contrast for the features in the cf_d faces. If the lowered contrast produced the aftereffect, we would expect to observe the motion-face-induced aftereffect on dot-free cartoon faces with reduced background luminance (47.6 cd/m^2) (Fig. 5b, cf_l). We ran the motion-face adaptation condition (mf- cf_l) and no-adaptation baseline condition (0- cf_l) in this luminance-matched case in exactly the same way that we ran the mf- cf_d and 0- cf_d conditions. The data (Fig. 6a–d, dashed and solid blue curves, and 6e, blue hatched bar) show no aftereffect, suggesting that luminance is not a critical factor. Furthermore, the slopes of the psychometric functions for the cf_l and cf_d conditions do not differ significantly (mean = 23.2 and 25.1, respectively; $p = 0.32$), suggesting that the different aftereffects cannot be attributed to different salencies of the test stimuli.

The second alternative explanation is that the added dots produced a motion aftereffect which then biased the perceived facial expressions. This is very unlikely because the added dots were outside the mouth curves (Fig. 5a) and thus had little or no spatial overlap with the moving dots inside the mouth of the adapting motion face. To rule out this possibility formally, we repeated the mf- cf_d condition but asked the subjects to report the direction of motion for the dots above the midpoint of the mouth curves of the test cartoon faces. The data for this condition (mf- cf_d -dir) are shown as green curves in Figure 6a–d. That the slopes of these

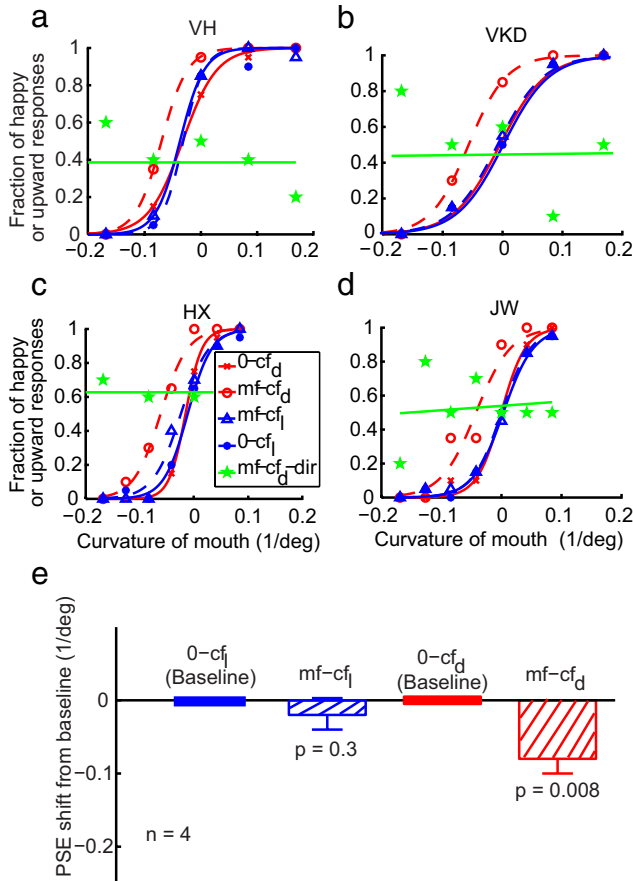


Figure 6. The role of background similarity in aftereffect transfer from second-order motion faces to first-order static cartoon faces. *a–d*, Psychometric functions from two naive subjects (VH, VKD) and two experimenters (HX, JW) under the following conditions: 0-cf_d, no adaptation and test on the dot-added cartoon faces (solid red); mf-cf_d, adaptation to the saddest motion face and test on the dot-added cartoon faces (dashed red); 0-cf_i, no adaptation and test on the background-luminance-matched cartoon faces (solid blue); mf-cf_i, adaptation to the saddest motion face and test on the background-luminance-matched cartoon faces (dashed blue); mf-cf_{d-dir}, same as mf-cf_d, but the subjects judged motion direction (up or down) above the midpoint of the mouth curves of the test cf_d faces (solid green). *e*, Summary of all four subjects' data.

curves are not significantly different from zero (mean, 0.014; $p = 0.62$) indicates that the subjects did not see motion aftereffects in this experimental condition. Indeed, the subjects complained that they could not perform this task because they could not see motion anywhere in the test faces.

The role of background similarity in determining aftereffect transfer between different types of first-order faces

We further investigated whether background similarity also contributes to aftereffect transfer between different types of first-order faces. We used real faces as in Figure 1*e* and static cartoon faces similar to those in Figure 1*d* but with a 30% pen width since there was no need to match the motion faces in this experiment. We first cross-adapted rf and cf stimuli without any background manipulations. There were four adaptation conditions rf-rf, cf-cf, rf-cf, and cf-rf, and two baseline conditions 0-rf and 0-cf. We found that cartoon-face adaptation generated a facial-expression aftereffect on the real faces (the cf-rf condition) but real-face adaptation failed to produce an aftereffect on the cartoon faces (the rf-cf condition). This is similar to the asymmetry between the motion and cartoon faces (Figs. 3 and 4), and we show only the summary data in Figure 7.

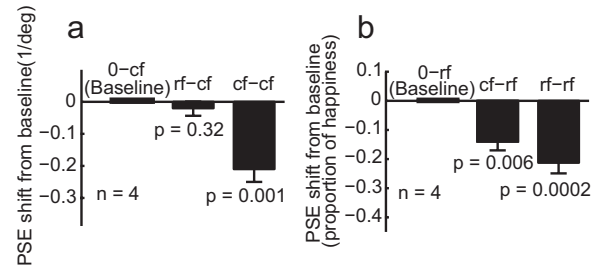


Figure 7. Cross-adaptation between first-order cartoon and real faces. Only the summary data from two naive subjects and two experimenters are shown. *a*, The test stimuli were always the cartoon faces. The adapting stimuli varied with conditions as follows: 0-cf, no-adaptation baseline; cf-cf, adaptation to the saddest cartoon face; rf-cf, adaptation to the saddest real face. *b*, The test stimuli were always the real faces. The adapting stimuli varied with conditions as follows: 0-rf, no-adaptation baseline; rf-rf, adaptation to the saddest real face; cf-rf, adaptation to the saddest cartoon face.

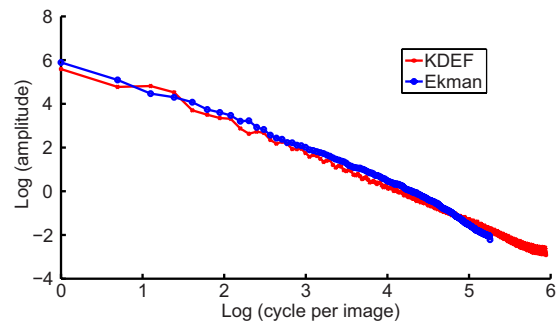


Figure 8. Correlation statistics of front-view real-face images measured by the Fourier amplitude spectrum of the KDEF database (red) and Ekman database (blue). The log–log plots show nearly straight lines, indicating a $1/f^k$ relationship between amplitude and spatial frequency. The KDEF curve extends to higher frequencies because of the larger image size. Curve fitting yields k values of 1.61 for the KDEF faces and 1.65 for Ekman faces, with an average of 1.63.

We then examined whether we can increase the aftereffect in the rf-cf condition by adding proper backgrounds to the test cartoon faces. One possibility would be to remove eyes, nose, and mouth from the adapting real face and use the remainder as the background for the test cartoon faces. However, the remainder might still carry some facial expression information or contain other features that could complicate the interpretation of the results. We therefore generated noise backgrounds based on the correlation statistics of real faces. For this purpose, we first computed the amplitude spectrum of real faces and found that it follows the $1/f^k$ law with $k = 1.63$ (Fig. 8), where k is the spatial frequency, in agreement with a recent study (Keil, 2008). Comparing with natural image statistics ($k = 1$) (Field, 1987; Ruderman, 1994), front-view faces have less energy in high-frequency range, presumably because of less abrupt occlusions.

We then repeated the rf-cf condition after adding $1/f^k$ noises to the test cartoon faces (denoted as the rf-cf_k conditions). We also run the corresponding baseline conditions 0-cf_k. In separate blocks with randomized order, we let $k = 0, 0.5, 1, 1.5, 2.0,$ and 2.5 . By histogram equalization, noises with different k values all had the same first-order distribution but different second-order correlation structures. (Note that first- and second-order image statistics should not be confused with first- and second-order stimuli.) We also included a uniform background whose luminance matches the mean of the other backgrounds; it is labeled as $k = \infty$. Samples of different backgrounds are shown in Figure 5*c*.

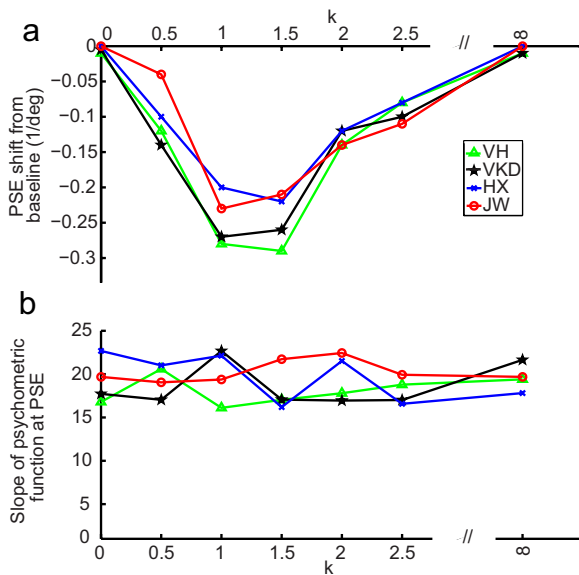


Figure 9. The role of background similarity in aftereffect transfer from first-order real faces to first-order static cartoon faces. The adapting stimulus was always the saddest real face. The test stimuli were cartoon face sets with $1/f^k$ noise backgrounds. *a*, The aftereffect, measured as the PSE shift from the corresponding baseline, as a function of k for four subjects. *b*, The slope (at PSE) of psychometric curves for the baseline conditions as a function of k .

The background in each trial was generated online. Consequently, although backgrounds for different trials in a block had identical statistics (same k), no sample was repeated. This ensures that any effect is not due to accidental “features” in a particular sample. Figure 9*a* shows the facial-expression aftereffect transfer from real-face adaptation to the test cartoon faces, measured as the PSE shift from the corresponding baseline condition, as a function of k . When the noise backgrounds had k values of 1 and 1.5, matching the correlation statistics of natural images and real-face images, the transfer was maximal. For k values outside this range, the transfer declines to 0. An ANOVA test indicates a significant dependence of the aftereffect on k ($p = 0.03$). Since all backgrounds (except the uniform one labeled as $k = \infty$) had the same first-order luminance distribution but different second-order correlation structures, Figure 9*a* suggests that the second-order statistics of the real and cartoon faces have to be similar to produce a sizable aftereffect.

An alternative interpretation of Figure 9*a* is that the test cartoon faces had different saliencies against different noise backgrounds even though they all had identical first-order luminance distributions, and differences in saliency led to different aftereffect transfer. To rule out this interpretation, we measured the slopes (at PSE) of the psychometric functions for different k values, and the results for the baseline conditions are shown in Figure 9*b*. An ANOVA test showed no significant dependence of the slope, and thus saliency, on k ($p = 0.41$). A similar lack of dependence on k was found for the slopes in the adaptation conditions ($p = 0.37$).

Discussion

We investigated the role of background statistics in aftereffect transfer between different classes of faces. We first generated a novel class of second-order faces in which local directions of motion define sad and happy facial expressions, without any static form cue. Reversing motion directions in these stimuli exchanges happy and sad expressions. We then examined interac-

tions between these second-order motion faces and first-order cartoon and real faces. Adaptation to a second-order motion face failed to produce a facial-expression aftereffect on the first-order real or cartoon faces, whereas adaptation to a first-order cartoon face, but not a real face, generated the aftereffect on the second-order motion faces, a difference that was not due to differences in contrast or mouth shape. These results agree with previous studies using simpler nonface stimuli, namely, that second-order adaptation typically does not transfer to first-order stimuli whereas first-order adaptation sometimes, but not always, transfers to the second-order stimuli (Nishida et al., 1997; Larsson et al., 2006; Ashida et al., 2007; Schofield et al., 2007). Such results have been widely interpreted as indicating separate processing for first- and second-order stimuli.

However, as mentioned in Introduction, cue-invariant cells with similar tuning to first- and second-order stimuli have been reported in many visual areas, raising the question of why aftereffect transfer is not more frequently observed, particularly from second- to first-order stimuli. We therefore searched for conditions that may facilitate transfer. We found that, surprisingly, the degree of similarity between the featureless backgrounds of adapting and test faces is a critical factor. Specifically, adaptation to a second-order motion face could be made to transfer to the first-order cartoon faces by simply adding static random dots to the latter making their backgrounds more similar to that of the adapting motion face (Fig. 6, the mf-cf_a condition). Moreover, background similarity also controlled transfer between different classes of first-order faces: Real-face adaptation transferred to static cartoon faces only when noise with the correlation statistics of real faces or natural images was added to the background of the cartoon faces (Fig. 9*a*). We ruled out contrast, motion-aftereffect, or saliency-based explanations for these findings. Our results have implications for mechanisms of adaptation, the interpretation of null aftereffects, and the representation of faces.

The background similarity effect in visual adaptation

The background similarity effect we described could be considered a new form of contingent aftereffect. The best-known contingent aftereffect was reported by McCollough (1965), who induced two different color aftereffects simultaneously, for two different orientations. This implies that a particular color aftereffect is observed only when the test and adapting stimuli have the same orientation. Similar contingent aftereffects have been reported for face adaptation (Rhodes et al., 2004; Yamashita et al., 2005; Ng et al., 2006; Little et al., 2008); in particular, facial-expression aftereffects are larger when the identities of adapting and test faces are shared (Fox and Barton, 2007). Our work extends these studies, showing the even greater importance of sharing gross statistical characteristics of otherwise featureless backgrounds.

The background similarity effect depends on both first- and second-order statistics. We originally demonstrated the effect by showing that transfer from motion faces with only first-order, white noise, random-dot backgrounds, to cartoon faces, occurred when similar, first-order, random dots were added to the latter. However, the backgrounds of real faces have second-order structure; substantial transfer from them to cartoon faces depended on matching this; simply matching the first-order distribution was insufficient. Of course, faces (and natural images in general) obviously contain higher-order structures. It would be interesting to test what aspects of higher-order statistics, notably those important for the determination of face identity (Fox and

Barton, 2007), could further enhance the background similarity effect.

At a neuronal level, contingent aftereffects likely arise from cells jointly selective to multiple aspects of stimuli. The background similarity effect, in particular, predicts joint selectivity to stimulus features and background statistics. Variation in firing rates coming from the background may therefore have resulted in an underestimation of the degree of invariance to features or cues defining the features, effectively comprising a novel nonclassical influence on visual responses. However, face-identity aftereffects are not influenced by differing facial expressions (Fox et al., 2008), showing that identity coding is invariant to expression-related statistics. How selectivity and invariance depends on tasks and levels of the visual hierarchy is an open question.

It thus becomes pressing to ask whether background manipulations can enhance aftereffect transfer between first- and second-order *low-level* stimuli such as orientation and translational motion. Cue-invariant responses to orientation and motion have been found in areas V1, V2, and MT (von der Heydt et al., 1984; Albright, 1992; Sheth et al., 1996). Computational models also assume that separate first- and second-order motion processing in V1 and V2, respectively, converges on MT (Wilson et al., 1992). However, it is possible that cue convergence is progressively stronger at higher-level areas coding more complex stimuli, with the fraction of cue-invariant cells in relatively low-level areas being too small (O'Keefe and Movshon, 1998) to have a perceptual impact. We are currently investigating this issue. Interestingly, the rare study that did find strong cross-adaptation aftereffects between first- and second-order orientations used stimuli of very similar backgrounds (Georgeson and Schofield, 2002).

At a functional level, our study suggests that statistical similarity between featureless backgrounds of stimuli can gate their temporal interactions. Although the statistical rationale for the adaptation effects that are observed is not yet completely clear (Schwartz et al., 2007; Xu et al., 2008), by restricting the scope of adaptation, the background similarity effect will avoid overgeneralization.

Adaptation treats the temporal context of stimuli. However, there are many similarities between the effects of spatial and temporal context (Clifford et al., 2000; Schwartz et al., 2007), as in the relationship between the tilt aftereffect and the tilt illusion. It is not known if the latter is gated by statistical similarity between the backgrounds of the stimuli involved. Crowding is another form of contextual spatial interaction and it is known to be stronger when the features in the target and flanks are more similar (Levi, 2008). However, it is also unknown if featureless backgrounds modulate the crowding effect.

Interpretation of null aftereffects

Our results suggest that if adaptation to A has no impact on B, then it is not necessarily true that A and B are processed separately. Our initial null findings might have tempted one to conclude that facial expressions in motion faces, real faces, and static cartoon faces are all processed separately, despite expression being a universal property. Our later finding of significant transfer for sufficiently similar backgrounds suggests that different cues do indeed converge in the course of creating facial-expression representation. The same issue may affect the design and interpretation of a range of future adaptation experiments. Nevertheless, that transfer depends on appropriate background manipulations suggests limitations in the extent of shared neural processing for the stimuli. In this sense, our results complement, rather than contradict, previous suggestions of separate processing of first- and second-order stimuli. A probable sce-

nario is that there are both separate and shared representations of different classes of stimuli. Previous studies emphasized the former whereas our method of background manipulation is sensitive enough to reveal the latter.

The notion of background similarity is symmetric: if A is similar to B, then B is similar to A; yet the interactions between different classes of stimuli are often asymmetric. An obvious explanation is that the outcome of a particular adaptation experiment depends on multiple factors including saliency and separate processing as well as background similarity.

Face representation

Our results suggest the possibility that background statistics are an integral component of facial-expression representation even when the backgrounds do not contain any information about facial expression. In other words, featureless backgrounds are not discarded after feature extraction. For example, when adapting to the saddest motion face, the context provided by the featureless stationary dots in the background may be encoded together with the dots that define the mouth shape (and facial expression). This integration could occur either at face areas or at lower-level areas feeding the face areas, and may be related to the holistic view of face representation (Tanaka and Farah, 2003).

The perceived expressions in our motion-face stimuli also suggest strong interactions between motion processing and face perception, functions that have been assigned to dorsal and ventral pathways, respectively (Ungerleider and Mishkin, 1982; Felleman and Van Essen, 1991). Although there have already been demonstrations of the interactions (Bassili, 1978; Berry, 1990; Knappmeyer et al., 2003; Farivar et al., 2009), our motion-face stimuli are unique in that the happy and sad expressions are solely defined by local *directions* of motion according to the motion mislocalization effect. Since reversing the directions exchanges the happy and sad expressions, our stimuli link face perception to directional processing, instead of general motion detection. Directional tuning and the associated mislocalization effect start as early as V1 (Hubel and Wiesel, 1968; Fu et al., 2004). Thus, our stimuli suggest that local, low-level directional processing contributes to face perception, supporting the notion of a strong local component in face representation, in addition to a holistic component (Xu et al., 2008).

In summary, we have shown that statistical similarity between featureless backgrounds of the adapting and test stimuli is a novel factor in determining aftereffect transfer between different classes of stimuli. This result has implications for the design and interpretation of future adaptation studies. We have also shown that facial expressions can be solely defined by local directions of motion gradient patterns, thus establishing a strong link between face and motion-direction processing. Our findings suggest that luminance- and motion-direction-defined forms, and local features and background statistics, converge in the representation of faces.

References

- Albright TD (1992) Form-cue invariant motion processing in primate visual cortex. *Science* 255:1141–1143.
- Ashida H, Lingnau A, Wall MB, Smith AT (2007) fMRI adaptation reveals separate mechanisms for first-order and second-order motion. *J Neurophysiol* 97:1319–1325.
- Atick JJ, Redlich AN (1990) Towards a theory of early visual processing. *Neural Comput* 2:308–320.
- Barlow H (1990) A theory about the functional role and synaptic mechanism of visual aftereffects. In: *Vision: coding and efficiency* (Blakemore C, ed). Cambridge, UK: Cambridge UP.

- Bassili JN (1978) Facial motion in the perception of faces and of emotional expression. *J Exp Psychol Hum Percept Perform* 4:373–379.
- Berry DS (1990) What can a moving face tell us? *J Pers Soc Psychol* 58:1004–1014.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. *Neuron* 26:695–702.
- Clifford CW, Wenderoth P, Spehar B (2000) A functional angle on some after-effects in cortical vision. *Proc R Soc Lond B Biol Sci* 267:1705–1710.
- Dan Y, Atick JJ, Reid RC (1996) Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *J Neurosci* 16:3351–3362.
- De Valois RL, De Valois KK (1991) Vernier acuity with stationary moving Gabors. *Vision Res* 31:1619–1626.
- Ekman P, Friesen W (1976) Pictures of facial affect. Palo Alto, CA: Consulting Psychologists.
- Ellemberg D, Allen HA, Hess RF (2004) Investigating local network interactions underlying first- and second-order processing. *Vision Res* 44:1787–1797.
- Farivar R, Blanke O, Chaudhuri A (2009) Dorsal–ventral integration in the recognition of motion-defined unfamiliar faces. *J Neurosci* 29:5336–5342.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.
- Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical-cells. *J Opt Soc Am A Opt Image Sci Vis* 4:2379–2394.
- Fox CJ, Barton JJ (2007) What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Res* 1127:80–89.
- Fox CJ, Oruç İ, Barton JJS (2008) It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *J Vis* 8:11.1–13.
- Fu YX, Shen Y, Gao H, Dan Y (2004) Asymmetry in visual cortical circuits underlying motion-induced perceptual mislocalization. *J Neurosci* 24:2165–2171.
- Geesaman BJ, Andersen RA (1996) The analysis of complex motion patterns by form/cue invariant MSTd neurons. *J Neurosci* 16:4716–4732.
- Georgeson MA, Schofield AJ (2002) Shading and texture: separate information channels with a common adaptation mechanism? *Spat Vis* 16:59–76.
- Gosselin F, Schyns PG (2001) Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Res* 41:2261–2271.
- Grill-Spector K, Kushnir T, Edelman S, Itzhak Y, Malach R (1998) Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21:191–202.
- Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195:215–243.
- Keil MS (2008) Does face image statistics predict a preferred spatial frequency for human face processing? *Proc R Soc Lond B Biol Sci* 275:2095–2100.
- Knappmeyer B, Thornton IM, Bühlhoff HH (2003) The use of facial motion and facial form during the processing of identity. *Vision Res* 43:1921–1936.
- Larsson J, Landy MS, Heeger DJ (2006) Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *J Neurophysiol* 95:862–881.
- Leopold DA, O'Toole AJ, Vetter T, Blanz V (2001) Prototype-referenced shape encoding revealed by high-level after effects. *Nat Neurosci* 4:89–94.
- Levi DM (2008) Crowding—an essential bottleneck for object recognition: a mini-review. *Vision Res* 48:635–654.
- Little AC, DeBruine LM, Jones BC, Waitt C (2008) Category contingent aftereffects for faces of different races, ages and species. *Cognition* 106:1537–1547.
- Lundqvist D, Flykt A, Öhman A (1998) The Karolinska Directed Emotional Faces (KDEF), CD-ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.
- McCullough C (1965) Color adaptation of edge-detectors in the human visual system. *Science* 149:1115–1116.
- Ng M, Ciaramitaro VM, Anstis S, Boynton GM, Fine I (2006) Selectivity for the configural cues that identify the gender, ethnicity, and identity of faces in human cortex. *Proc Natl Acad Sci U S A* 103:19552–19557.
- Nishida S, Ledgeway T, Edwards M (1997) Dual multiple-scale processing for motion in the human visual system. *Vision Res* 37:2685–2698.
- O'Keefe LP, Movshon JA (1998) Processing of first- and second-order motion signals by neurons in area MT of the macaque monkey. *Vis Neurosci* 15:305–317.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Ramachandran VS, Anstis SM (1990) Illusory displacement of equiluminous kinetic edges. *Perception* 19:611–616.
- Rhodes G, Jeffery L, Watson TL, Jaquet E, Winkler C, Clifford CW (2004) Orientation-contingent face aftereffects and implications for face-coding mechanisms. *Curr Biol* 14:2119–2123.
- Ruderman DL (1994) The statistics of natural images. *Netw Comput Neural Syst* 5:517–548.
- Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260:995–997.
- Schofield AJ, Ledgeway T, Hutchinson CV (2007) Asymmetric transfer of the dynamic motion aftereffect between first- and second-order cues and among different second-order cues. *J Vis* 7(8):1.
- Schwartz O, Hsu A, Dayan P (2007) Space and time in visual context. *Nat Rev Neurosci* 8:522–535.
- Sheth BR, Sharma J, Rao SC, Sur M (1996) Orientation maps of subjective contours in visual cortex. *Science* 274:2110–2115.
- Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- Smith S, Clifford CW, Wenderoth P (2001) Interaction between first- and second-order orientation channels revealed by the tilt illusion: psychophysics and computational modelling. *Vision Res* 41:1057–1071.
- Tanaka JW, Farah MJ (2003) The holistic face representation. In: Perception of faces, objects and scenes. Analytic and holistic processes (Peterson MA, Rhodes G, eds), pp 53–74. New York: Oxford UP.
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: Analysis of visual behavior (Ingle DJ, Goodale MA, Mansfield RJW, eds), pp 549–586. Cambridge, MA: MIT.
- von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. *Science* 224:1260–1262.
- Wainwright MJ (1999) Visual adaptation as optimal information transmission. *Vision Res* 39:3960–3974.
- Webster MA, Kaping D, Mizokami Y, Duhamel P (2004) Adaptation to natural facial categories. *Nature* 428:557–561.
- Wilson HR, Ferrera VP, Yo C (1992) A psychophysically motivated model for 2-dimensional motion perception. *Vis Neurosci* 9:79–97.
- Xu H, Dayan P, Lipkin RM, Qian N (2008) Adaptation across the cortical hierarchy: low-level curve adaptation affects high-level facial-expression judgments. *J Neurosci* 28:3374–3383.
- Yamashita JA, Hardy JL, De Valois KK, Webster MA (2005) Stimulus selectivity of figural aftereffects for faces. *J Exp Psychol Hum Percept Perform* 31:420–437.
- Zhaoping L (2006) Theoretical understanding of the early visual processes by data compression and data selection. *Netw Comput Neural Syst* 17:301–334.