Bidirectional Gender Face Aftereffects: Evidence Against Normative Facial Coding

Sophie L. Cronin, Morgan L. Spence, Paul A. Miller and Derek H. Arnold
Perception Laboratory, School of Psychology, The University of Queensland, QLD, Australia

Abstract
Facial appearance can be altered, not just by restyling but also by sensory processes. Exposure to a female face can, for instance, make subsequent faces look more masculine than they would otherwise. Two explanations exist. According to one, exposure to a female face renormalizes face perception, making that female and all other faces look more masculine as a consequence—a *unidirectional* effect. According to that explanation, exposure to a male face would have the opposite unidirectional effect. Another suggestion is that face gender is subject to contrastive aftereffects. These should make some faces look more masculine than the adaptor and other faces more feminine—a *bidirectional* effect. Here, we show that face gender aftereffects are bidirectional, as predicted by the latter hypothesis. Images of real faces rated as more and less masculine than adaptors at baseline tended to look even more and less masculine than adaptors post adaptation. This suggests that, rather than mental representations of all faces being recalibrated to better reflect the prevailing statistics of the environment, mental operations exaggerate differences between successive faces, and this can impact facial gender perception.

Facial appearance is malleable. After exposure to a face (often referred to as an adaptation period), the shape (Webster & MacLin, 1999), identity (Leopold, O'Toole, Vetter, & Blanz, 2001), gender (Afraz & Cavanagh, 2008), race (Amihai, Deonell, & Bentin, 2011), emotional expression (Fox, Oruç, & Barton, 2008), and attractiveness (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003) of subsequent faces can all appear altered.

Two explanations exist. According to one, face aftereffects (AEs) happen because appearance is determined by the contrast between a given face and an encoded facial norm

Corresponding author:
Derek H. Arnold, School of Psychology, McElwain Building, St Lucia, Brisbane, QLD 4072, Australia.
Email: d.arnold@psy.uq.edu.au

(Anderson & Wilson, 2005; Giese & Leopold, 2005; Leopold et al., 2001; Rhodes & Jeffery, 2006; Rhodes & Leopold, 2011; Susilo, McKone, & Edwards, 2010; Tsao & Freiwald, 2006; Valentine, 1991; Webster & MacLeod, 2011). The norm is thought to reflect the average face encountered in all circumstances throughout life. As the facial norm is an average of male and female faces, it should appear androgynous. Hypothetically, each face we encounter acts as an "adaptor" which updates the norm, bringing it closer to the adaptors' *physical* characteristics. If the adaptor was a male face, the norm would be shifted toward masculine physical characteristics, and *all* faces would look more feminine as a consequence. Previously masculine looking faces would look more feminine as there would be less difference between these and the new norm, whereas previously feminine looking faces would look more feminine as there would now be a greater physical difference between these and the updated norm (see Pond et al., 2013).

Computationally, it is thought that a dimensional norm (e.g., facial gender) corresponds with the physical input that elicits equal responses from two oppositely tuned channels, and that norm updating is achieved by differentially changing the responsiveness of these two channels. This scenario is often referred to as a two-channel norm-based opponent coding scheme (Anderson & Wilson, 2005; Giese & Leopold, 2005; Leopold et al., 2001; Rhodes & Jeffery, 2006; Rhodes & Leopold, 2011; Susilo et al., 2010; Tsao & Freiwald, 2006; Webster & MacLeod, 2011; see Figure 1(a)).

Another explanation is that facial appearance is determined by a multichannel coding scheme, in a similar fashion to orientation perception (Barlow & Hill, 1963; Hubel & Wiesel, 1962). In these schemes, perception is determined by the average stimulus value signaled by active channels. Each channel "neuron" can be conceptualized as voting for a given stimulus value, in proportion to that channel's firing rate. Post adaptation, units responsive to the adaptor become less responsive, and so the averaging process becomes biased toward values offset from the adaptor. These biases are signaled by channels that are now relatively more responsive. Importantly, such biases are *bidirectional*, toward values offset from the adaptor in either direction. Hence, if face gender was subject to multichannel coding, as has been suggested (Storrs & Arnold, 2012, 2015; Zhao, Seriès, Hancock, & Bednar, 2011), faces that had appeared more and less masculine than the adaptor at baseline will tend to appear even more and less masculine post adaptation (see Figure 1(b)). This can be described as a contrastive AE, as this coding scheme tends to exaggerate differences between the adaptor and distinct subsequent inputs.

While norm-based and multichannel coding predict different patterns of perceptual AE (unidirectional and bidirectional), it can be difficult to diagnose data as resulting from either scheme. The major reason for this is that both schemes make similar predictions concerning category boundaries, and this is what is typically measured in facial adaptation studies (Afraz & Cavanagh, 2008; Pond et al., 2013; Zhao et al., 2011). Both schemes, for instance, predict that the face gender category boundary (the point at which people begin to categorize faces as male more often than female) will shift toward faces that are physically more masculine post male face adaptation. Norm-based coding predicts this because the putative "norm" will have shifted toward the adapted value, creating a discrepancy between this and faces about the old normative value, which had seemed androgynous, but will now seem more feminine due to the changed reference frame. Multichannel coding predicts this because adaptation will exaggerate differences between the adapted and other stimulus values, in this instance biasing perception of what had seemed androgynous away from the masculine adaptor, toward a feminine appearance (see Figure 1(b)).

One attempt to diagnose AEs as arising from norm-based or multichannel coding focused on post adaptation sensitivity changes (Lawson, Clifford, & Calder, 2011). To attribute these
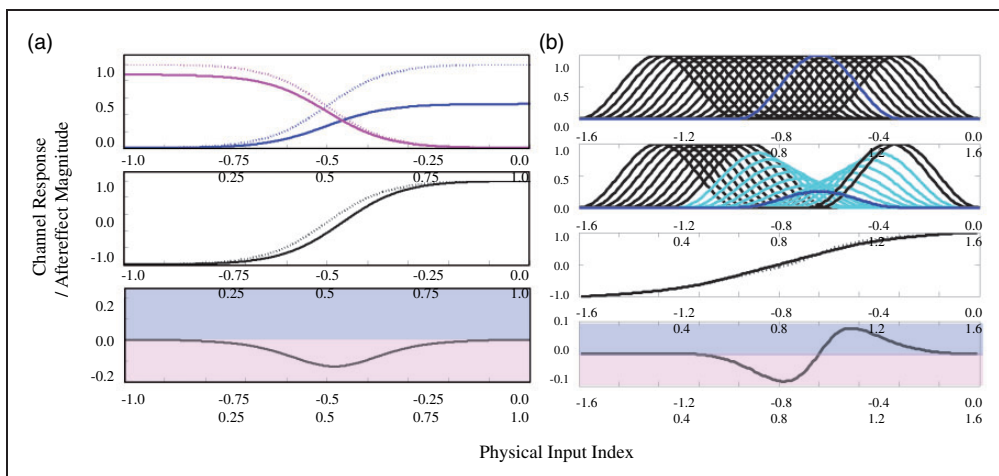
Figure 1. Depiction of two coding schemes, and the AE patterns they predict. (a) In a two-channel norm-based opponent code, two channels (top row—dotted lines unadapted, bold adapted) are maximally responsive to opposite values on a dimension, here to female (-1 on the physical input index scale) and male ($+1$). When active, these channels can be thought of as "voting" for their preferred value in proportion to their response. Male face adaptation is simulated by reducing the male channel responsiveness relative to the female channel. A discriminant function, of masculinity estimates as a function of physical input, can be determined by expressing male "votes" in proportion to the total vote for each input value (middle row). Post adaptation the discriminant function is shifted toward more masculine physical inputs (bold line), relative to baseline (dotted line). This produces a unidirectional AE, with all affected faces looking more feminine (bottom row). (b) In a multichannel coding scheme, channels are tuned to different physical input values (Row 1). Each channel can be thought of as voting for a specific masculinity estimate. Adaptation to a somewhat male face can be simulated by reducing the responsiveness of a subset of channels, maximally for the channel tuned to the adapted value (Row 2). Masculinity estimates can be determined by summing all votes, with each "vote" expressed in proportion to all votes (Row 3). This generates contrastive AEs, with some inputs encoded as less masculine post adaptation, and others as more masculine (Row 4).

to sensory coding, it is necessary to presume decisional criteria used in both phases of the experiment are either fixed or equally variable (see Yarrow, Jahn, Durant, & Arnold, 2011). In the first instance, any change in the pattern of behavioral response can unambiguously be attributed to coding changes, as decisional factors are assumed to be fixed, but this is not possible if one accepts that decisional criteria might be more stable in the adaptation condition than in the baseline. The constant repetition of a particular exemplar (or alternation of two exemplars), which characterizes face adaptation protocols, could encourage participants to adopt the adaptor(s) as a reference, and judge other images relative to it. An unadapted baseline condition, however, might be marked by less rigid criteria, as there is no repeated reference against which to judge other test images. Hence, responses in an adaptation condition could be more systematic, in addition to being more biased, because of decisional factors, and this could be misconstrued as an improvement in sensitivity due to coding changes. There is thus some ambiguity concerning whether changes in decisional sensitivity can be attributed to sensory coding, or to the adoption of more or less rigidly defined decisional criteria.

   Other attempts to diagnose AEs as arising from a given coding scheme have focused on the *magnitude* of category boundary shifts post adaptation. Both schemes predict larger shifts of

the category boundary, at least initially, as the adaptor becomes a more extreme example of a masculine or feminine face. Multichannel coding predicts that such AEs should eventually diminish, as adaptors cease to influence neurons that are at least somewhat responsive to both the adaptor and to nearly androgynous test faces. However, if support was found for this (see Zhao et al., 2011), norm-based coding proponents could argue that the relevant adaptors had been too extreme, and had ceased to look human, thus explaining the reduction in AE magnitude (see McKone, Jeffery, Boeing, Clifford, & Rhodes, 2014; Pond et al., 2013 and Figure 1(a)). If AE magnitudes seem to scale with adaptor distance from the category boundary (see Pond et al., 2013), multichannel coding proponents could argue that insufficiently extreme adaptors had been sampled. Such evidence is therefore ambiguous and inconclusive.

There are further problems inherent in trying to use the magnitude and spread of an AE, as measured at the category boundary, as a coding scheme diagnostic. To predict how an AEs' magnitude should change according to a multichannel coding scheme, one needs to know how far the facial dimension extends. Does the facial gender dimension, for instance, encompass unnatural looking ogrish-masculine and elfin-feminine faces, or are such unnatural (but still recognizable) faces encoded by different populations of neurons? Is the dimension in question uniformly sampled by neurons responding to different subsets of faces (evenly spaced along the dimension), or is the dimension nonuniformly sampled? What range of different faces will a given face-tuned neuron respond to (i.e., what is the neurons' bandwidth tuning)? Finally, where precisely do your experimental stimuli fall within the facial dimension? This last question is impossible to know without answering the first question posed here. Assumed answers to all these questions will determine how rapidly a multichannel code predicts an AE should increase and decline as differences between a particular adaptor and set of test images increase. The magnitude and spread of AEs predicted by opponent coding are subject to similar considerations, such as the degree to which adaptation changes channel responsiveness. Our point, which we wish to make forcefully, is that it is very difficult to make firm predictions about how the magnitude of a face AE should change for a set of inputs that are increasingly different to an adaptor, as experimenters lack sufficient insight into a host of potentially influential factors.

A more definitive test is possible, but it necessitates an assessment of the appearance of inputs falling *both* above and below an adaptor in terms of the relevant dimension. Specifically, for face gender, we want to know if adaptation to any given face exaggerates the masculinity of more masculine faces, while also exaggerating the femininity of more feminine faces. Importantly, this clear diagnostic is predicted by a multichannel coding scheme but not by a norm-based opponent code (see Figure 1). Note that this approach (assessing the appearance of faces to either side of an adaptor in terms of the facial dimension) is markedly different to the most common methodology used in this context, which involves estimating the positioning of a natural category boundary (i.e., the point at which faces switch from being categorized as predominantly male to female, or from being Caucasian to Asian, or from being normally shaped to distorted; see Afraz & Cavanagh, 2008; Amihai et al., 2011; Webster & MacLin, 1999).

As face gender adaptation has retinally localized effects (see Afraz & Cavanagh, 2008), we can assess the appearance of faces (independent of their position relative to a category boundary) by comparing images presented in different locations. As both opponent and multichannel coding can predict similar changes to the positioning of a natural category boundary (Afraz & Cavanagh, 2008; Pond et al., 2013; Zhao et al., 2011), we have chosen instead to focus on a more diagnostic test—whether face AEs are unidirectional (*as predicted by opponent coding*) or bidirectional (*as predicted by multichannel coding*).

Experiment 1

Methods

Thirty volunteers participated (aged 17–28, $M = 20.3$, $SD = 2.8$, 20 female). All were naïve as to the experimental purpose and had normal or corrected-to-normal visual acuity. This number of participants was predetermined based on numbers of participants included in other, conceptually similar, studies.

Experimental stimuli were generated using MATLAB R2012b software (MathWorks, Natick, MA) in conjunction with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and were presented on either a 19″ Samsung SyncMaster 950SL, a 19″ Samsung SyncMaster 950p+, or a 19″ Dell Trinitron monitor, all set to a 1024×768 pixel resolution and a refresh rate of 60 Hz. Stimuli were viewed from 57 cm with the head placed in a chin rest. Responses were recorded via mouse button presses.

Twenty images (10 male and 10 female) of famous faces were chosen. These depicted front-facing celebrities adopting neutral expressions. Images were chosen on the basis that they had no facial hair, hair obscuring the face, or visible accessories (such as jewelry or spectacles). Images were cropped around the face line and converted to grayscale.

Exemplar male and female faces were generated by averaging the 10 celebrity facial images of either gender using the face mixer function in Abrosoft FantaMorph. An exemplar androgynous face was generated using the same process to average all 20 images of both genders (see Figure 2). Images were resized to subtend 2.2 degrees of visual angle (dva) in diameter at the retina. The average luminance of all images was equated, first by using the SHINE Toolbox for MATLAB (Willenbockel et al., 2010) and then via an additional MATLAB calibration routine. The outline of faces was obscured via a gray oval-shaped ring, resulting in an oval-shaped visible face without hair or ears (see Figure 2). These were shown against a black background, so the average luminance and the average luminance-contrast of all images relative to background were equated.

Procedure

Participants also completed a block of unadapted trials. On each trial, two test images were presented (for 0.5 seconds). One was centered 2.2 dva to the left of a central red fixation point, the other 2.2 dva to the right. Each subtended 2 dva in diameter. After each test
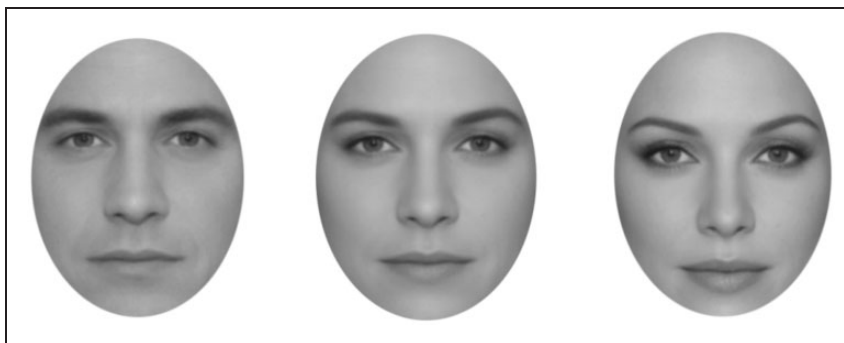

Figure 2. Exemplar male (left), androgynous (middle), and female (right) faces. All of these were included as test images, and male and female exemplars were used as adapting images.

presentation, participants indicated which face had looked more masculine by pressing either the left or right mouse button. During a block of trials, each of the 23 test images (20 celebrity images, 3 averaged exemplars) were used as standards and compared with the remaining 22 test images. A block of trials consisted of 506 individual trials, all presented in random order with a 0.5-second pause between each trial. Participants were instructed to maintain central fixation throughout testing.

Participants then completed an adaptation block of trials. Adaptation trials were similar to baseline trials, with the exception that test presentations were preceded by a 3.75-second adaptation presentation, and then by a 1-second interstimulus-interval (ISI). In both adaptation and baseline blocks of trials, responses were unhurried. Adapting images were larger (subtending 2.6 dva in diameter at the retina) versions of the male and female exemplars and were centered on the same locations on the display as test images. Half of the participants adapted to the male exemplar on the right and the female on the left, other participants adapted to a reversed contingency. A 1-second ISI followed each adaptation presentation (see Figure 3). Midway through a block of trials, a screen was presented instructing participants to rest and then to resume testing when ready.

Results

Our paradigm necessitates a slight modification to the qualitative AE pattern predicted by two-channel norm-based opponent coding (see Figure 1). During adaptation blocks of trials, each test image will be contrasted with all other test images in the other test location (see Figure 3), and these images will have been subject to *opposite* adaptation. This will *exaggerate*
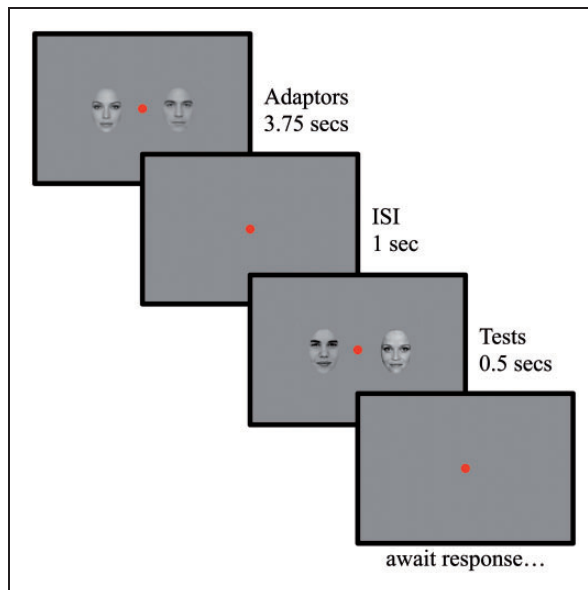

Figure 3. Graphic depicting a trial sequence during an adaptation block of trials. Here, a scenario is depicted wherein the participant has adapted to the female exemplar on the left and the male on the right. Following adaptation, there was an interstimulus-interval, followed by a test presentation. The participant then indicated which test image had looked more masculine.

AEs predicted by opponent coding, by an amount equal to the *average* impact of the opposite adaptor on other tests (see Figure 4(a)). This does not, however, complicate our core opponent coding prediction—that face gender AEs should be *unidirectional*.

The use of opposite adaptors could also modulate multichannel coding predictions. If other tests are evenly distributed about the opposite adaptor, no modification to the shape of the predicted AE function is necessary, as equal numbers of other tests should seem more masculine and more feminine, producing no net bias. If, however, other tests are unevenly distributed about the opposite adaptor, the AE function could be vertically shifted, by an amount equal to the *average* impact of the opposite adaptor on other tests. In Figure 4(b), we have depicted two extreme scenarios, wherein other tests are subject, on average, to half the maximal AE induced by the opposite adaptor. Note that this would shift the AE function inflexion point (where no AE is predicted) *away* from the adaptor's position along the encoded dimension. Note also that this would introduce an asymmetry, with differently sized maximal positive and negative AEs. Once more, however, our core multichannel coding prediction holds. This scheme will tend to encourage simultaneous *oppositely signed* AEs for tests positioned above and below the adaptor in terms of the encoded stimulus dimension. Moreover, while multichannel coding tends to predict this bidirectionality, *opponent coding cannot*—so bidirectionality remains a strong diagnostic even if precise AE predictions must be modulated.
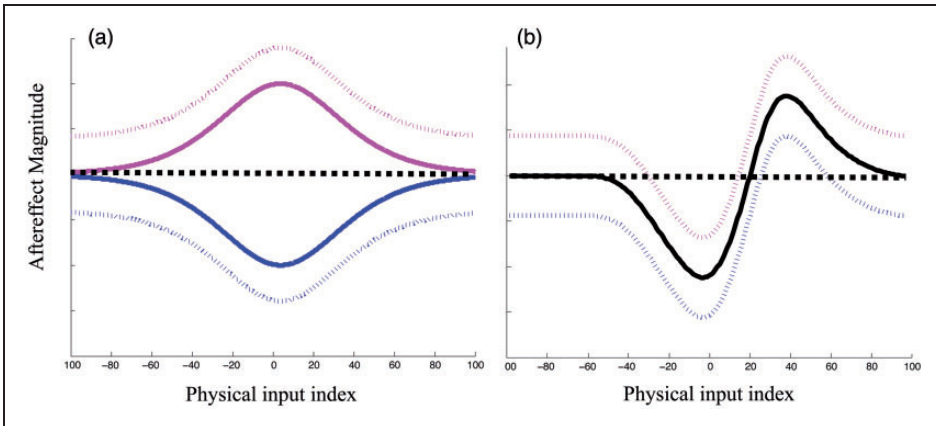
Figure 4. Modified predictions for Experiment 1. In this experiment, the appearance of oppositely adapted tests was contrasted. As each test was compared with all other (oppositely adapted) tests, opponent coding (a) predicts AE magnitudes will be exaggerated, by the average impact of the opposite adaptor on all other tests. If adaptation makes a test look more masculine, opposite adaptation will make other tests look more feminine, increasing the apparent masculinity of the test image in comparison. Opposite adaptation could also modulate population-coding predictions (b). If other tests are evenly distributed about the opposite adaptor, no modification is necessary (half other tests would look less masculine, half more masculine, producing no net bias). If, however, other tests are disproportionately positioned about the opposite adaptor, AE scores will be modulated by the average impact of the opposite adaptor. Here, two opposite extreme scenarios are depicted, wherein other test images are subject to half the maximum AE in either direction. Note that relative to an unmodulated AE function, this would induce a lateral shift (such that the function inflexion point, where no AE occurs, shifts away from the adaptor). Note also the predicted asymmetry in terms of maximal AE for tests below and above the adaptor.

To facilitate a test for a bidirectional AE pattern, as opposed to a unidirectional pattern, we can contrast AEs for tests that are *exaggerated* relative to the adaptor—test faces that are rated as more feminine than a female adaptor in that adaptation condition, and test faces rated as more masculine than a male adaptor in that adaptation condition. Multichannel coding predicts that perceptual differences, between these tests and their adaptors, will be exaggerated, so faces rated as more feminine than a female adaptor at baseline should seem even more feminine post female adaptation (*producing negative AE scores*), and faces rated more masculine than a male adaptor should look even more masculine post male adaptation (*producing positive AE scores*, see Figure 5(a)). Opponent coding predicts the opposite in each case—that after adapting to a female face, all test faces (including those rated as more feminine than the adaptor at baseline) should look more masculine, producing *positive AE scores*. Conversely, after adapting to a male face, all test faces should seem more feminine (including those rated as more masculine at baseline), producing *negative AE scores* (see Figure 5(b)). The two coding schemes therefore predict *oppositely* signed contrast values when we subtract exaggerated male adapted AE scores from exaggerated female adapted AE scores.

Contrast values from comparing exaggerated test image AE scores can be compared with contrast values from comparing muted test image AE scores. Each coding scheme predicts same-signed AEs in these circumstances (see Figure 5(b))—faces rated more feminine than a female adaptor at baseline should look more masculine post adaptation (*producing positive AE scores*), and faces rated less masculine than a male adaptor should look more feminine post adaptation (*producing negative AE scores*). The critical test for our hypothesis thus reduces to whether comparisons of these two sets of AEs will produce contrast values of the same (*predicted by opponent coding*) or opposite (*predicted by multichannel coding*) sign.

Each block of trials provided a masculinity rating (MR) for each test image. These reflect the proportion of trials in which that particular test image had been rated as *more* masculine relative to *all* other test images presented in the other test location. A MR of 0 would indicate that the test had *never* been judged as more masculine than other test images, whereas a MR of 1 would indicate that it had *always* been judged as more masculine.

Individual AE scores were calculated for each test image in each adaptor condition (tests in same position as the male adaptor, or in the same position as the female adaptor). This was done by subtracting baseline MRs from adapted MRs (i.e., AE score = Adapted MR − Baseline MR). Negative AE scores signify a *reduction* in MRs post adaptation relative to baseline, positive AEs an *increase* (see Figure 6).

As it would be harder to detect a tendency for our most masculine looking test images to look even more masculine relative to other tests, and vice versa for very feminine faces; AE scores were transformed by expressing raw AE scores as a proportion of the *possible* change in the observed AE direction (positive or negative from baseline) given the baseline MR for that image for that participant. For example, an image with a baseline MR of 0.4 and an adapted MR of 0.2 would have a transformed AE score of −0.5 (as it had undergone half the possible change given the AE direction). The same image with an adapted MR of 0.6 would have a transformed AE score of +0.33.

Our core predictions are that face gender AEs should either be unidirectional (according to an opponent coding account) or bidirectional (according to a multichannel coding account). To assess this, we can contrast differences between AE scores for tests that are *exaggerated* relative to their adaptor (so tests that look more feminine than the female adaptor, and tests that look more masculine than the male adaptor) and compare these to contrast values from comparing AE scores for tests that are *muted* relative to their adaptor (so tests that look less feminine than the female adaptor, and tests that look less masculine than the male adaptor).
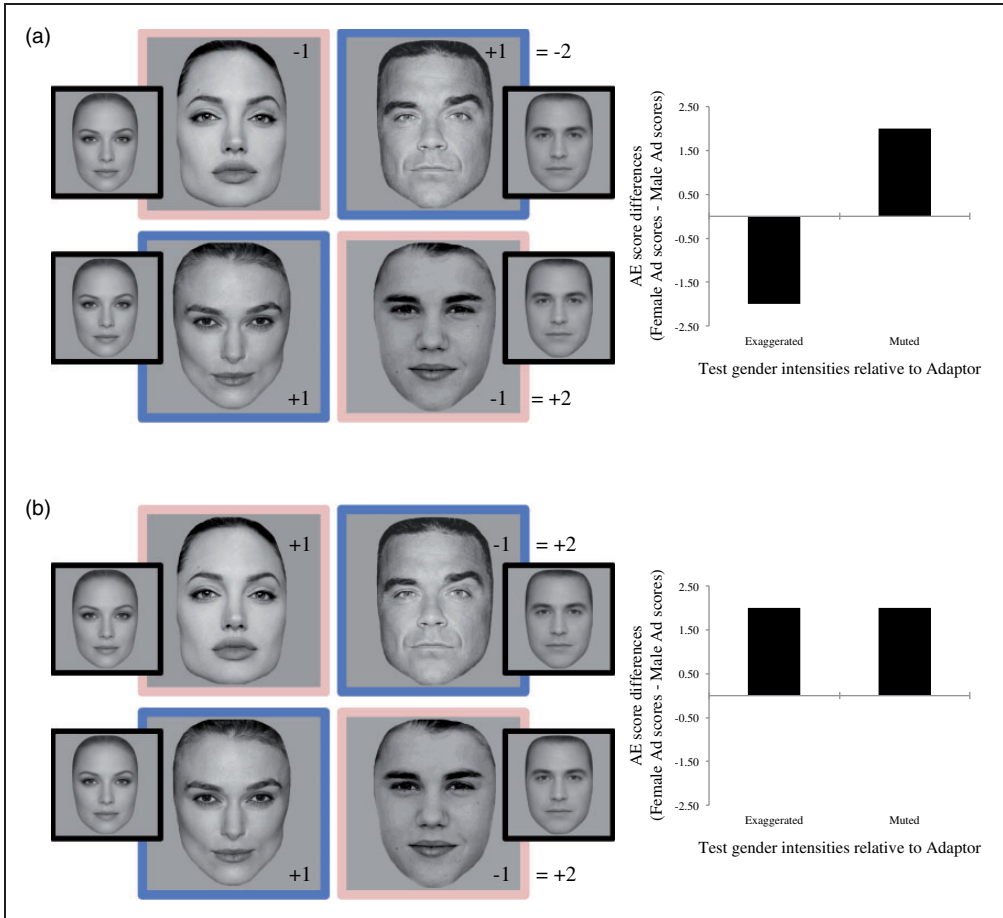
Figure 5. Graphic depicting predictions concerning the sign of face gender AE score contrasts, from multichannel coding (a) and opponent coding (b). Blue frames signify that the test image was rated as more masculine than the relevant adaptor at baseline, pink that the test image was rated as less masculine. Relevant adaptors are depicted as smaller images, with black frames, adjacent to test images. In each panel, test images that are exaggerated in terms of the adapted gender are shown above, whereas those that are muted are shown below. Multichannel coding predicts that AE directions will be contingent on whether an input is positioned below (-1) or above ($+1$) the relevant adaptor in terms of the category dimension. So multichannel coding (a) predicts a negative contrast value when comparing AE scores for images exaggerated in terms of gender relative to adaptors at baseline—images rated more feminine than the female adaptor (-1) and images rated more masculine than the male adaptor ($+1$, so -1 minus $+1 = -2$). Opponent coding instead predicts AE directions will be contingent on the adaptor, so opposite biases should ensue from adapting to a female ($+1$) or to a male (-1) adaptor. For the same comparison as above, opponent coding (b) predicts a positive AE contrast value (i.e., $+1$ minus $-1 = +2$). Note that the two coding schemes predict the same outcome for images muted in terms of gender relative to adaptors—images rated less feminine than the female adaptor ($+1$), and images rated less masculine than the male adaptor ($+1$, so $+1$ minus $-1 = +2$). Average AE magnitudes will impact on the scale of this prediction, but will not impact on the sign of these model predictions.
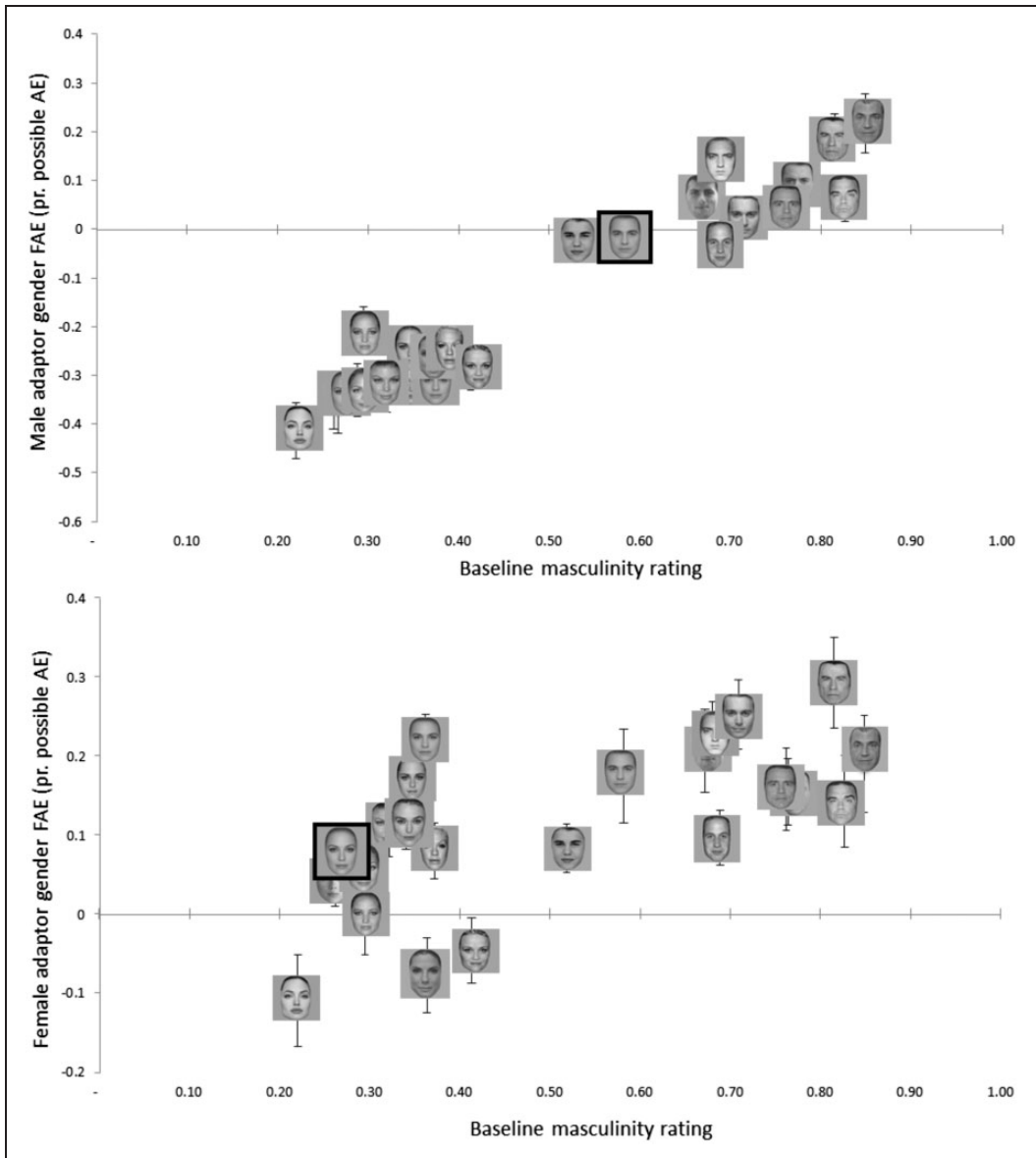
Figure 6. AE scores, averaged across participants, for test images as a function of baseline MRs, averaged across participants. Data are plotted separately for male (above) and for the female (below) adaptation conditions. Miniature test images are superimposed over relevant data points. Adapting images are highlighted by bold black frames. When visible beyond the limits of test images, error bars show ± 1 *SEM*.

AE scores for exaggerated and muted test images are shown in Figure 6(a), for both the female and male adaptor conditions.

All AEs were in the direction predicted by our multichannel coding account. Exaggerated female adapted AEs had, on average, a negative sign (-0.35, $SD = 0.265$), whereas exaggerated male adapted AEs had, on average, a positive sign (0.09, $SD = 0.032$). Subtracting individual exaggerated male AE scores from individual exaggerated female AE

scores resulted in a positive contrast value (0.125, $SD = 0.384$; single sample $t(29) = 1.797$, $p = .0415$, single tailed), as predicted by multichannel coding, and contrary to the predictions of opponent coding (see Figure 5). Muted female adapted AEs had, on average, a positive sign (0.124, $SD = 0.16$), whereas muted male adapted AEs had, on average, a negative sign (-0.259, $SD = 0.202$). Subtracting individual muted male AE scores from individual muted female AE scores resulted in a negative contrast value (-0.383, $SD = 0.346$; single sample $t(29) = 6.083$, $p < .0001$, single tailed). Note that we have used single-tailed t tests here, as we are testing for a predicted AE sign.

Discussion

The pattern of results in Experiment 1 was precisely as predicted by our multichannel coding account, and contrary to that predicted by opponent coding. However, a number of reasonable concerns can be raised. First, in this experiment, we used images of famous faces as stimuli, and this could have conceivably distorted our results. It is unclear how this might have encouraged a pattern of results reminiscent of the predictions of either of the two coding schemes in question, but we nonetheless we repeated the experiment using nonfamous faces.

Another consideration is the conceptually unnecessary complication of having included an opposite adaptor. While this did not impact on our core qualitative predictions, it could reasonably be regarded as unnecessary. So in Experiment 2, participants adapted to a single, male or female, exemplar positioned to the left or right of fixation, with comparison images presented on the other side of fixation.

Experiment 2

Methods

All details concerning Experiment 2 were as for Experiment 1, with the following exceptions. Thirty-five volunteers participated (aged 17–41, $M = 21.9$, $SD = 5.5$, 30 female). All were naïve as to the experimental purpose. Half adapted to a male image and half to a female image—a between-groups design. Only one adaptation image was used in each case, presented to the left or right of fixation, counterbalanced for both groups. Images of famous faces from Experiment 1 were replaced by pictures of noncelebrities sourced from Google images. Image calibrations were as for Experiment 1.

Results

Initial stages of data analysis were as for Experiment 1, up to and including the calculation of AE scores. These are depicted, for each adaptation group (male or female), in Figure 7.

Given the between-groups design of Experiment 2, we cannot calculate individual contrast values for tests that were *exaggerated* in terms of gender relative to the adaptor in each adaptor condition. To assess our core prediction, we shall therefore conduct two single sample t tests, testing for a robust face AE for *exaggerated* tests in each group. Recall that the two coding schemes in question predict *oppositely* signed AEs for these two conditions—multichannel coding predicts *negative* AEs for exaggerated female tests and positive AEs for exaggerated male tests, whereas opponent coding predicts the reverse contingency (see Figure 5). Analyses of these data supported multichannel coding predictions, with a negatively signed AE for exaggerated female tests ($M = -0.11$, $SD = 0.23$; single sample $t(17) = 3.09$, $p = .0035$, single tailed) and a positively signed AE
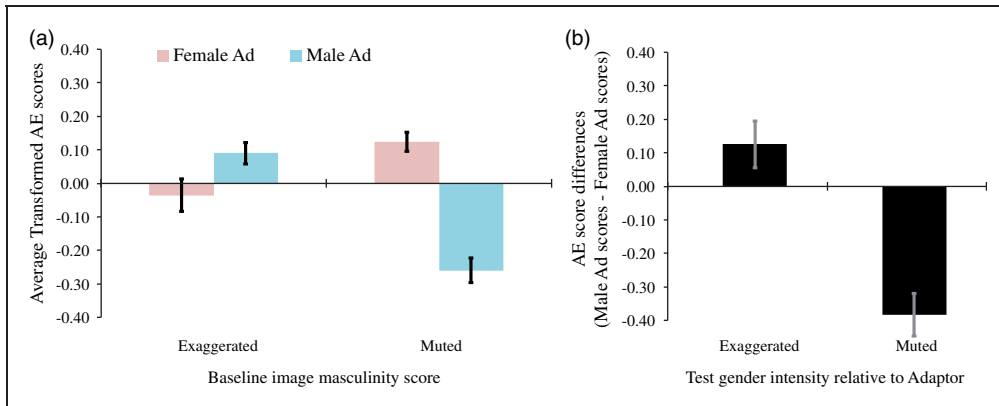
Figure 7. (a) AE scores for test images that were exaggerated in terms of gender relative to their adaptor at baseline (more masculine than the male adaptor, and more feminine than the female adaptor) and for test images that were muted in terms of gender relative to their adaptor (less masculine than the male adaptor and less feminine than the female adaptor). Data are shown for the female adaptor (magenta) and for the male adaptor (cyan). (b) Contrast scores from comparing AEs for exaggerated female adapted and exaggerated male adapted test images. In all cases, error bars represent ±1 *SEM*.

for exaggerated male tests ($M = 0.16$, $SD = 0.32$; single sample $t(16) = 2.09$, $p = .0263$, single tailed—see Figures 8 and 9).

Discussion

AE directions in Experiments 1 and 2 were as predicted by multichannel coding and were inconsistent with opponent coding predictions. While this repeated pattern of results seems compelling, these data could be criticized as imprecise. The AE functions (see Figures 6 and 8) did not seem precisely defined, with a pivot point neatly falling about the adapting image. A substantial degree of measurement noise is to be expected in this context, and so we regard *any* robust region of oppositely signed AEs for exaggerated, relative to muted, test images in either adapting condition as strong disproof that opponent coding solely drives these AEs. Regardless, one factor potentially contributing to measurement noise is our use of relatively extreme adaptors in terms of the masculinity dimension. A cleaner pattern of results might ensue if we instead have people adapt to a single adaptor, positioned centrally within the stimulus dimension. This would avoid complications arising from having opposite adaptors (as per Experiment 1) and avoid the necessity of a between-groups design (Experiment 2). In Experiment 3, we therefore had a single group adapt to a single androgynous exemplar image.

Experiment 3

Methods

All details concerning Experiment 3 were as for Experiment 2, with the following exceptions. Thirty-eight volunteers participated (aged 17–34, $M = 19.1$, $SD = 3.1$, 28 female). All were naïve as to the experimental purpose. All participants adapted to a single image depicting an androgynous exemplar face (see Figure 2). This was presented to the left or right of fixation, counterbalanced across participants. All other test images were as for Experiment 2.
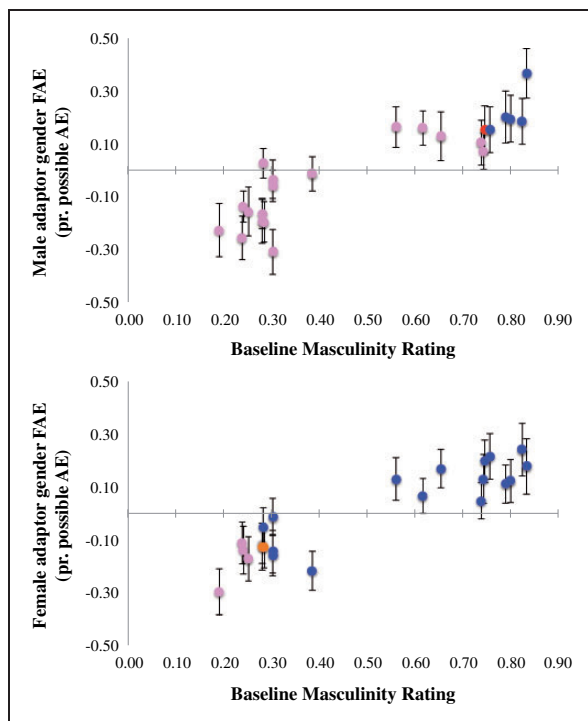
Figure 8. AE scores, averaged across participants, for test images as a function of baseline MRs, also averaged across participants. Data are plotted separately for the male (above) and for the female (below) adaptation groups. Data points for images rated as more masculine than the relevant adaptor at baseline are colored blue, data points for images rated as more feminine than the relevant adaptor are magenta. Adaptor data points are colored red. Error bars show ±1 *SEM*.
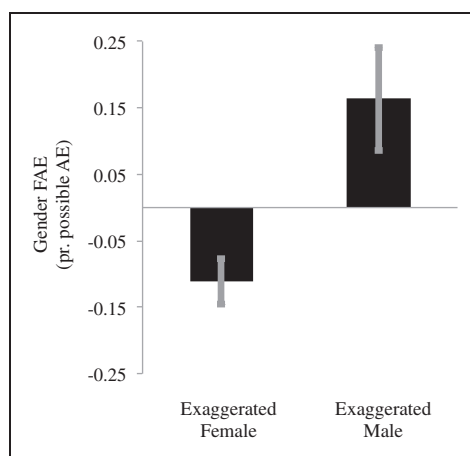
Figure 9. AE scores for test images that were exaggerated in terms of gender relative to their adaptor at baseline (more masculine than the male adaptor and more feminine than the female adaptor). Error bars represent ±1 *SEM*.

Results

AE scores, as a function of baseline MRs, are depicted in Figure 10. Perusal of this figure reveals a clear bidirectional AE pattern, as predicted by multichannel coding, but inconsistent with opponent coding predictions.

To assess our core prediction, we conducted two single sample t tests, testing for oppositely signed FAEs for tests regarded as more masculine and as more feminine at baseline than the adaptor. Of the two coding schemes, this is predicted by multichannel coding (see Figure 5), whereas opponent coding either predicts a single direction of AE or if the androgynous image we have chosen as an adaptor equally excites the hypothesized two-opponent channels—no AE at all. Analyses of these data support multichannel coding predictions, with a negatively signed AE for images rated as more feminine than the adaptor at baseline ($M = -0.14$, $SD = 0.07$; single sample $t(37) = 5.13$, $p < .0001$, single tailed) and a positively signed AE for images rated as masculine than the adaptor at baseline ($M = 0.24$, $SD = 0.06$; single sample $t(37) = 6.13$, $p < .0001$, single tailed—see Figures 8 and 11).

General Discussion

Our data show that face gender adaptation can exaggerate masculinity differences. Images rated as more feminine than adaptors at baseline tend to look even more feminine post adaptation, and images rated more masculine at baseline tend to look even more masculine post adaptation (see Figures 5 to 7). These data are consistent with contrastive AEs, as predicted by multichannel coding, but are inconsistent with the unidirectional AEs predicted by norm-based opponent coding (see Figure 1).

Our data agree with results recently reported in a conceptually similar study, contrasting the appearance of tests presented in differently adapted locations. That study examined tilt,
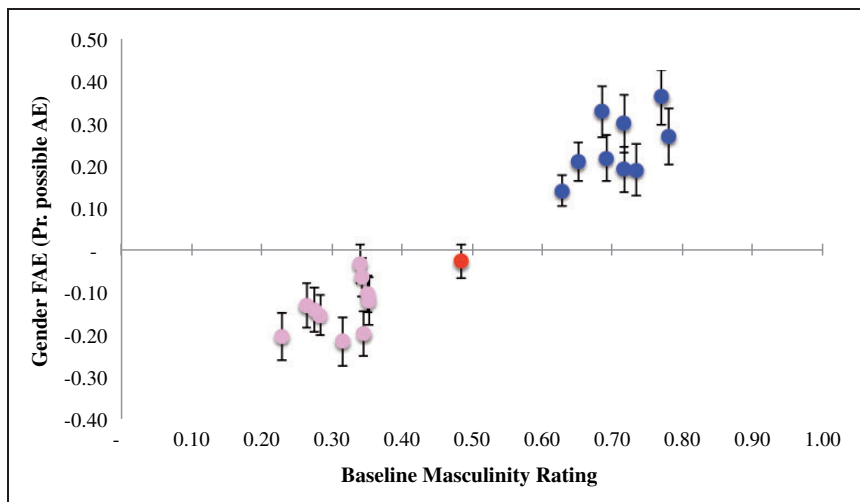
Figure 10. AE scores, averaged across participants, for test images as a function of baseline MRs, also averaged across participants. Data points for images rated as more masculine than the androgynous adaptor at baseline are colored blue, data points for images rated as more feminine than the androgynous adaptor are magenta. The adaptor's data point is colored red. Error bars show ±1 *SEM*.
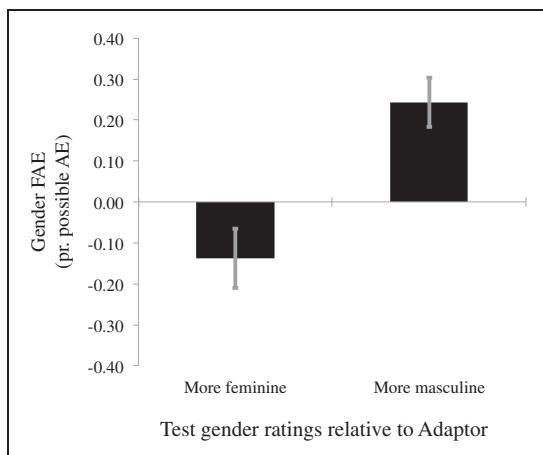
Figure 11. AE scores for test images rated as more feminine and as more masculine, on average across participants, than an androgynous exemplar image used as an adaptor. Error bars represent ± 1 *SEM*.

facial gender, and facial identity AEs. In all cases, AEs were mitigated when one of the two test images was also a (*scaled*) adaptor. This mitigation tends to be predicted by multichannel coding but is inconsistent with opponent coding predictions (see Storrs & Arnold, 2015). Here, we tested a stronger prediction of multichannel coding—that AEs should be bidirectional. This cannot be predicted by a two-channel opponent code, as these predict unidirectional AEs. Our data are in accord with the former prediction and at odds with the latter.

Test images in this study were intentionally smaller ($\sim$70%) than adaptors, so our data cannot readily be ascribed to operations that depend on adaptors and tests being tightly colocalized in retinal coordinates (Afraz & Cavanagh, 2008, 2009; Zhao & Chubb, 2001). Nor can our data be attributed to some test images, depicting celebrities, being recognized. First, this was unlikely due to the brief (0.5 seconds) peripheral test presentations. Second, any such tendency should have stabilized MRs, whereas the key feature of our data is that facial appearance was malleable, and finally—we obtained qualitatively matched results in two subsequent experiments using nonfamous faces. In all experiments, masculinity and femininity was exaggerated, depending on whether the test had been rated as more or less masculine than the relevant, unfamiliar, adaptor at baseline.

It is worth noting that in Experiment 1, the exemplar male image was rated as more effeminate than all but one of the celebrity male faces (a youthful Justin Bieber), which had been averaged to generate it (also see Pond et al., 2013). This is perhaps unsurprising, as averaging multiple facial images results in a smoothed symmetrical image, free of skin blemishes, which may be at odds with most people's expectations of a male face. This observation is pertinent, as several other studies have used similar methods to generate exemplar "male" images and have then generated an array of test images linearly morphing between this and a similarly generated female exemplar (Afraz & Cavanagh, 2009; Rhodes & Jeffery, 2006; Rhodes et al., 2004). Our data suggest that such studies are examining variance ranging from a somewhat androgynous to a very feminine-looking face and are unlikely to encompass masculine-looking faces. This strains credibility concerning norm-based coding interpretations of the data resulting from such investigations. Even when a study examined extrapolations beyond the natural limits of gender facial appearance, due to a generation process based on image averaging with initial end-points suggested by male

and female "exemplars," the testing range would have been offset, biased toward physical images that appeared feminine (see Zhao et al., 2011).

Conceptual Context—"Face Space"

The nature of the computations underlying face AEs are of interest as they speak to the nature of the computations underlying face perception. It has been suggested that face perception relies on a mapping that encapsulates the variance between different human faces (see Valentine, 1991). The arrangement of this mapping, however, remains mysterious. According to one view, facial information is mapped about a norm, with faces described in terms of how they vary from the norm (Rhodes, Brennan, & Carey, 1987; Rhodes & Tremewan, 1994; Valentine, 1991). This is known as a norm-based facial code. According to another view, mapping of facial information is more diffuse, with faces described in terms of how similar, or dissimilar, they are to various previously encountered faces (Byatt & Rhodes, 1998; Robert, 1999; Valentine, 1991; Wallis, Siebeck, Swann, Blanz, & Bülthoff, 2008). This is known as exemplar-based facial coding.

Some have taken face AE data as evidence for norm-based facial coding (Leopold, Bondar, & Giese, 2006; Rhodes & Jeffery, 2006; Susilo et al., 2010). The idea is that the norm is updated via adaptation, thereby changing the appearance of all subsequent faces because these are encoded in terms of how they vary from the updated norm. In essence, according to this view, mental representations are updated to reflect the prevailing statistics of the environment (Rhodes et al., 2003). It is important for such interpretations that AEs be unidirectional—a norm cannot simultaneously be updated in opposite directions. An exemplar-based coding scheme, however, is consistent with multichannel coding and can predict contrastive (bidirectional) AEs. According to this view, adaptation might serve to enhance sensitivity to change, by exaggerating physical differences between the adaptor and subsequent inputs (Calder, Jenkins, Cassel, & Clifford, 2008; Clifford, Ma Wyatt, Arnold, Smith, & Wenderoth, 2001). Our data argue in favor of the latter exemplar-based coding strategy for face gender.

We believe there were preexisting reasons to be sceptical about the norm-based coding explanation of face AEs. Given that an individual's face space is hypothetically constructed from a lifetime of experiences, it would be surprising if extremely brief (e.g., 5 seconds, Leopold et al., 2001) exposures to an adaptor could dramatically shift the implicit norm, but this would seem necessary according to this explanation. Rapid neural adaptation resulting in bidirectional contrastive AEs can, however, happen (see Kohn, 2007 for a review) . Further, it is possible to induce different AEs simultaneously, contingent on the position, orientation, and the gender of faces depicted in test images. For example, opposite configural distortions can be induced in upright and inverted images, such that upright test images seem to have more contracted features, and inverted images expanded features (Rhodes et al., 2004). If contemplated purely within the confines of opponent coding, this would seem to require that there are at least two somewhat distinct face spaces, one for upright and another for inverted faces (Rhodes et al., 2004). In fact, simultaneous AEs have been demonstrated for different races (Jaquet, Rhodes, & Hayward, 2007), spatial positions (Afraz & Cavanagh, 2008), and genders (Little, DeBruine, Jones, & Waitt, 2008; Jaquet & Rhodes, 2008), so contingent adaptation could be, and has been, taken to suggest not one face space but multiple somewhat independent spaces for faces with different orientations, genders, races, and retinal locations. While possible, these suggestions do not seem parsimonious, and all these data could equally, or perhaps be better, explained in terms of multichannel coding interactions that exaggerate differences between adaptors and distinct visual images presented in similar circumstances.

While we are sceptical about an opponent coding of facial attributes, we would not like to create the impression that opponent coding, or multichannel coding, are conceptually implausible. Both schemes successfully describe patterns of AE in different contexts. Color adaptation, for instance, is well described by opponent coding (Hurvich & Jameson, 1957; Webster & Leonard, 2008), whereas spatial frequency and orientation AEs are well described by multichannel coding (Blakemore & Nachmias, 1971; for a review, see Suzuki, 2005). Our point is that both schemes are conceptually plausible and could well be used to encode information about faces. Our data, however, argue against an opponent encoding of facial gender.

Caveats

We have championed using a combination of spatial adaptation and a comparative methodology to ascertain whether AEs are unidirectional (consistent with a two-channel opponent code) or bidirectional (consistent with a multichannel code). We consider this necessary, as the two coding schemes make very similar predictions about category boundary shifts (which is what is typically measured in face AE studies).

On the basis of our data, we can state with confidence that face gender AEs are not *solely* driven by opponent coding. This scheme cannot promote a bidirectional AE, and we have found strong evidence for one in this context. This leaves open the possibility that there might be an additional influence of opponent coding. If, for instance, the effects of opponent coding were universal, impacting all faces regardless of location, our spatial comparison task could not have measured it—we could only measure spatially specific AEs. Note, however, that while this possibility exists, the suggestion lacks parsimony, as it is not immediately apparent that one needs to assume this additional process to explain empirical observations. Also, evidence suggests any spatially nonspecific face AE is weak in comparison to retinally mapped AEs (see Afraz & Cavanagh, 2008, 2009), and it might not easily be differentiated from experience-based decisional biases (see Yarrow et al., 2011).

Conclusion

Overall, we would argue that in this context, too much emphasis has been placed on the results of binary categorization tasks (Leopold et al., 2006; Rhodes & Jeffery, 2006; Susilo et al., 2010). These can identify the positioning of a natural category boundary—the point at which people cease to primarily categorize an input as one thing and begin to categorize it as another. Binary categorization tasks cannot, however, diagnose an AE as unidirectional or bidirectional, and it is this that is most clearly diagnostic of a norm-based versus multichannel coding strategy. The comparative methodology we have used permits the appearance of stimuli offset in both directions from an adaptor to be assessed. Using this approach, we obtained data showing that the gender face AE is bidirectional, as predicted by multichannel coding but in contradiction to the predictions of norm-based opponent coding. Similar studies should be conducted to assess if other facial dimensions are subject to similar coding strategies.

Declaration of Conflicting Interests

Funding

References

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

Afraz, S. R., & Cavanagh, P. (2008). Retinotopy of the face aftereffect. *Vision Research*, *48*, 42–54. doi:10.1016/j.visres.2007.10.028

Amihai, I., Deouell, L., & Bentin, S. (2011). Conscious awareness is necessary for processing race and gender information from faces. *Consciousness and Cognition*, *20*, 269–279. doi:10.1016/j.concog.2010.08.004

Anderson, N. D., & Wilson, H. R. (2005). The nature of synthetic face adaptation. *Vision Research*, *45*, 1815–1828. doi:10.1016/j.visres.2005.01.012

Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects. *Nature*, *200*, 1345–1347. doi:10.1038/2001345a0

Blakemore, C., & Nachmias, J. (1971). The orientation specificity of two visual after-effects. *The Journal of Physiology*, *213*, 157–174. doi:10.1113/jphysiol.1971.sp009374

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Byatt, G., & Rhodes, G. (1998). Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research*, *38*, 2455–2468. doi:10.1016/S0042-6989(97)00469-0

Calder, A. J., Jenkins, R., Cassel, A., & Clifford, C. W. G. (2008). Visual representation of eye gaze is coded by a nonopponent multichannel system. *Journal of Experimental Psychology: General*, *137*, 244–261. doi:10.1037/0096-3445.137.2.244

Clifford, C. W. G., Ma Wyatt, A., Arnold, D. H., Smith, S. T., & Wenderoth, P. (2001). Orthogonal adaptation improves orientation discrimination. *Vision Research*, *41*, 151–159. doi:10.1016/S0042-6989(00)00248-0

Fox, C. J., Oruç, I., & Barton, J. J. (2008). It doesn't matter how you feel. *The facial identity aftereffect is invariant to changes in facial expression*. Journal of Vision, *8*, 11–13. doi:10.1167/8.3.11

Giese, M. A., & Leopold, D. A. (2005). Physiologically inspired neural model for the encoding of face spaces. *Neurocomputing*, *65*, 93–101. doi:10.1016/j.neucom.2004.10.060

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*, 106.

Hurvich, L. M., & Jameson, D. (1957). An opponent process theory of colour vision. *Psychological Review*, *64*, 384–404.

Jaquet, E., & Rhodes, G. (2008). Face aftereffects indicate dissociable, but not distinct, coding of male and female faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 101–112. doi:10.1037/0096-1523.34.1.101

Jaquet, E., Rhodes, G., & Hayward, W. G. (2007). Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences. *The Quarterly Journal of Experimental Psychology*, *60*, 1457–1467. doi:10.1080/17470210701467870

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, *97*, 3155–3164. doi:10.1152/jn.00086.2007

Lawson, R. P., Clifford, C. W. G., & Calder, A. J. (2011). A real head turner: Horizontal and vertical head directions are multichannel coded. *Journal of Vision*, *11*, 17. doi:10.1167/11.9.17

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, *442*, 572–575. doi:10.1038/nature04951

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. doi:10.1038/82947

Little, A. C., DeBruine, L. M., Jones, B. C., & Waitt, C. (2008). Category contingent aftereffects for faces of different races, ages and species. *Cognition*, *106*, 1537–1547. doi:10.1016/j.cognition.2007.06.008

McKone, E., Jeffery, L., Boeing, A., Clifford, C. W., & Rhodes, G. (2014). Face identity aftereffects increase monotonically with adaptor extremity over, but not beyond, the range of natural faces. *Vision Research*, *98*, 1–13. doi:10.1016/j.visres.2014.01.007

Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. doi:10.1163/156856897X00366

Pond, S., Kloth, N., McKone, E., Jeffery, L., Irons, J., & Rhodes, G. (2013). Aftereffects support opponent coding of face gender. *Journal of vision*, *13*, 16. doi:10.1167/13.14.16

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology*, *19*, 473–497. doi:10.1016/0010-0285(87)90016-8

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, *46*, 2977–2987. doi:10.1016/j.visres.2006.03.002

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science*, *14*, 558–566. doi:10.1046/j.0956-7976.2003.psci_1465.x

Rhodes, G., Jeffery, L., Watson, T. L., Jaquet, E., Winkler, C., & Clifford, C. W. G. (2004). Orientation-contingent face aftereffects and implications for face-coding mechanisms. *Current Biology*, *14*, 2119–2123. doi:10.1016/j.cub.2004.11.053

Rhodes, G., & Leopold, D. A. (2011). Adaptive norm-based coding of face identity. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford handbook of face perception* (pp. 263–286). Oxford, England: Oxford University Press.

Rhodes, G., & Tremewan, T. (1994). Understanding face recognition: Caricature effects, inversion, and the homogeneity problem. *Visual Cognition*, *1*, 275–311. doi:10.1080/13506289408402303

Robert, M. B. L. (1999). A unified account of the effects of caricaturing faces. *Visual Cognition*, *6*, 1–42. doi:10.1080/713756800

Storrs, K. R., & Arnold, D. H. (2012). Not all face aftereffects are equal. *Vision Research*, *64*, 7–16. doi:10.1016/j.visres.2012.04.020

Storrs, K. R., & Arnold, D. H. (2015). Face aftereffects involve local repulsion, not renormalisation. *Journal of Vision*, *15*, 1doi:10.1167/15.8.1

Susilo, T., McKone, E., & Edwards, M. (2010). Solving the upside-down puzzle: Why do upright and inverted face aftereffects look alike? *Journal of Vision*, *10*, 1–16. doi:10.1167/10.13.1

Suzuki, S. (2005). High-level pattern encoding revealed by brief shape aftereffects. In C. Clifford, & G. Rhodes (Eds.), *Fitting the mind to the world: Adaptation and aftereffects in high level vision* (pp. 135–172). Oxford, England: Oxford University Press.

Tsao, D. Y., & Freiwald, W. A. (2006). What's so special about the average face? *Trends in Cognitive Sciences*, *10*, 391–393. doi:10.1016/j.tics.2006.07.009

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, *43*, 161–204. doi:10.1080/14640749108400966

Wallis, G., Siebeck, U. E., Swann, K., Blanz, V., & Bülthoff, H. H. (2008). The prototype effect revisited: Evidence for an abstract feature model of face recognition. *Journal of Vision*, *8*, 1–15. doi:10.1167/8.3.20

Webster, M. A., & Leonard, D. (2008). Adaptation and perceptual norms in colour vision. *Journal of the Optical Society of America A*, *25*, 2817–2825. doi:10.1364/JOSAA.25.002817

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society of London*, *366*, 1702–1725. doi: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. doi:10.3758/BF03212974

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*, 671–684. doi:10.3758/BRM.42.3.671

Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, *20*, 1518–1531. doi:10.1016/j.concog.2011.07.003

Zhao, C., Seriès, P., Hancock, P. J. B., & Bednar, J. A. (2011). Similar neural adaptation mechanisms underlying face gender and tilt aftereffects. *Vision Research*, *51*, 2021–2030. doi:10.1016/j.visres.2011.07.014

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