



Selectivity of face aftereffects for expressions and anti-expressions

Igor Juricevic¹ and Michael A. Webster^{2*}

¹ Department of Psychology, Indiana University, South Bend, IN, USA

² Department of Psychology, University of Nevada, Reno, NV, USA

Edited by:

Peter James Hills, Anglia Ruskin University, UK

Reviewed by:

Gyula Kovács, Budapest University of Technology, Hungary

Fang Fang, Peking University, China

Harold Hill, University of Wollongong, Australia

*Correspondence:

Michael A. Webster, University of Nevada, Reno, NV 89557, USA.
e-mail: mwebster@unr.edu

Adapting to a facial expression can alter the perceived expression of subsequently viewed faces. However, it remains unclear whether this adaptation affects each expression independently or transfers from one expression to another, and whether this transfer impedes or enhances responses to a different expression. To test for these interactions, we probed the basic expressions of anger, fear, happiness, sadness, surprise, and disgust, adapting to one expression and then testing on all six. Each expression was varied in strength by morphing it with a common neutral facial expression. Observers determined the threshold level required to correctly identify each expression, before or after adapting to a face with a neutral or intense expression. The adaptation was strongly selective for the adapting category; responses to the adapting expression were reduced, while other categories showed little consistent evidence of either suppression or facilitation. In a second experiment we instead compared adaptation to each expression and its anti-expression. The latter are defined by the physically complementary facial configuration, yet appear much more ambiguous as expressions. In this case, for most expressions the opposing faces produced aftereffects of opposite sign in the perceived expression. These biases suggest that the adaptation acts in part by shifting the perceived neutral point for the facial configuration. This is consistent with the pattern of renormalization suggested for adaptation to other facial attributes, and thus may reflect a generic level of configural coding. However, for most categories aftereffects were stronger for expressions than anti-expressions, pointing to the possible influence of an additional component of the adaptation at sites that explicitly represent facial expressions. At either level our results are consistent with other recent work in suggesting that the six expressions are defined by dimensions that are largely independently normalized by adaptation, possibly because the facial configurations conveying different expressions vary in independent ways.

Keywords: adaptation, aftereffects, face perception, facial expressions

INTRODUCTION

Facial expressions are important stimuli for signaling our emotional states and thus are critically involved in many social functions. Most humans are consequently adept at recognizing them, and failures in recognition are symptomatic of serious cognitive and neurological impairments (Calder et al., 2001). The human face displays an enormous variety of expressions, including a set of six basic expressions of emotion that correspond to happiness, anger, sadness, fear, disgust, and surprise (Ekman, 1992). The facial configurations signaling these states reflect highly stereotyped action patterns (Ekman and Friesen, 1978) and are to a large extent (though not completely, e.g., Russell, 1994) common across cultures, suggesting that they are primarily innate and universal.

An actively investigated question is how information about expressions is encoded in the visual system, and whether different expressions are represented by common or distinct pathways. Functionally, expressions convey information about affect, and it remains unclear whether basic emotions are independent or represent complementary or related states. For example, circumplex models of affect hold that different emotions are polar opposites

(Plutchik, 2001) or represent differences in a smaller number of underlying dimensions such as valence or arousal (Russell, 1980; Posner et al., 2005). The perception of expressions involves changeable aspects of the face and is thought to involve cortical areas which differ from the areas that are primarily responsible for invariant aspects of the face such as identity. Specifically, the Superior Temporal Sulcus has been implicated in expression recognition while identity recognition has instead pointed to the importance of a distinct network including the Fusiform Gyrus (Kanwisher et al., 1997; Haxby et al., 2000; Rossion et al., 2003). Many different neural structures appear dedicated to generating and processing the basic expressions of emotion (Adolphs et al., 1994, 1995, 1996; Morris et al., 1998; Sprengelmeyer et al., 1998; Kesler/West et al., 2001; Said et al., 2011), and thus the relationships between these different categories are complex and still unresolved. On the one hand, diverse evidence from studies of disease, lesions, and neuroimaging have revealed partially shared pathways for some expressions. Yet on the other hand, the same approaches have also provided widespread evidence for selective impairments and activation patterns for the perception of different expressions,

arguing strongly against the possibility that all expressions are encoded as dimensions of a common single representation (Calder et al., 2001).

In this study we examined the visual coding of facial expressions by measuring how the perception of expressions changes with adaptation. Viewing a stimulus can lead to large aftereffects in the appearance of subsequent stimuli. These adaptation effects have been widely used as a tool for probing the visual coding of stimulus features like color, motion, or orientation (Webster, 2011), and recently a number of studies have used adaptation to examine the processing of facial configurations (Webster and MacLeod, 2011). The appearance of a face can be strongly biased by prior adaptation. For example Webster and MacLin (1999) showed that adapting to a distorted face (e.g., one in which the features are expanded) induces a negative aftereffect in the appearance of the original face (e.g., so that the face appears too contracted). Similar negative aftereffects have been found for many of the dimensions that characterize natural variations in faces, including individual identity (Leopold et al., 2001) and facial categories such as gender and ethnicity (Webster et al., 2004).

Several previous studies have demonstrated that perceived expression can be biased by prior exposure to a face with a different expression (Russell and Fehr, 1987; Hsu and Young, 2004; Webster et al., 2004; Fox and Barton, 2007; Furl et al., 2007a,b; Benton and Burgess, 2008; Ellamil et al., 2008; Rutherford et al., 2008; Skinner and Benton, 2010; Cook et al., 2011; Pell and Richards, 2011). These experiments have thus shown that – like other aspects of face perception – the perception of facial expressions is highly adaptable. Importantly, this work has also suggested that the adaptation depends in part on the high-level configural properties of the face, and not simply on low-level properties such as the local features, nor on conceptual properties such as the conveyed emotion (Fox and Barton, 2007; Butler et al., 2008; Rutherford et al., 2008). (However, low-level features can also contribute; Xu et al., 2008.) Thus the adaptation to facial expressions appears to tap into visual pathways that may at least partly mediate the visual recognition of expressions, and may therefore provide a method for exploring how information is organized within these pathways.

A number of these studies have previously explored the interaction between different expressions. For example, Hsu and Young (2004) found that adapting to fearful, happy, or sad expressions produced selective losses in sensitivity to the adapted emotion, but also showed some facilitation across the expressions. Rutherford et al. (2008) instead asked observers to label the expression of a face with a neutral expression after adapting to each basic expression. They also observed asymmetric interactions where negative expressions were similar in inducing more responses than the neutral face appeared happy, while adapting to the happy expression caused the neutral test to specifically be judged as more sad. Pell and Richards (2011) further found an asymmetric relationship between the aftereffects for anger, fear, and disgust and argued from these that these expressions were encoded in partially overlapping representations. In contrast, Skinner and Benton (2010) recently reported that adaptation to faces with anti-expressions (formed by morphing each basic facial expression through an average expression and thus toward a face image with the opposite facial configuration) produced highly selective changes in the

ratings for each expression. For example, a face with the opposite expression of happy selectively increased the probability of judging the average face as happy. More recently Cook et al. (2011) instead explored adaptation effects along the principal axes of variation in natural expressive poses of a face (so that the axes were not tied to the canonical expressions). They showed that adaptation to positive or negative excursions along the first or second principal axis led to opposing aftereffects along that axis, but not to the (second or first) orthogonal axis.

The results of these studies thus differ in the extent to which adaptation to one expression might influence the perception of other expressions. In turn, this has implications for understanding the extent to which the visual encoding of different expressions might be separable (at least at the coding levels affected by the adaptation). In this study we sought to further explore this question by measuring how adaptation to each basic expression affected the sensitivity to different expressions. In particular, we assessed the changes in the recognition of each expression relative to a face with a neutral expression. The stimulus spaces explored by Skinner and Benton (2010) and Cook et al. (2011) – which have provided the strongest evidence for norm-based representation of expression – were instead anchored by the average expression in their samples. This has the advantage that the reference is defined by the stimulus distribution, but the disadvantage that this average could itself appear non-neutral and in particular could convey a possible expression. An average of two expressions can appear strongly biased after adapting. For example, viewing a happy or angry face biases the perceived expression of an intermediate morph between the two expressions toward the unadapted face (Webster et al., 2004). We took advantage of the fact that for expressions there is a “psychologically neutral” face pose defined by the neutral expression, and then asked how the canonical expressions defined as trajectories relative to this reference interacted in the adaptation. To address this question, we first conducted an experiment that examined how adaptation to one expression affected the recognition of the same or different expressions. In a second experiment, we instead asked how this recognition was affected when observers were adapted to one of the basic expressions or to the corresponding “anti-expression” representing the opposite configural change in the face.

MATERIALS AND METHODS

SUBJECTS

Observers included author IJ and 17 additional observers who participated either voluntarily or for partial course credit and who were naïve with respect to the aims of the study. A total of 12 subjects were tested in the first experiment and 7 in the second, with IJ tested in both. All had normal or corrected-to-normal vision. Participation was with informed consent and all experiments followed protocols approved by the university’s Institutional Review Board.

STIMULI

We used two different sets of stimuli for the two experiments – one which allowed us to assess the adaptation effects for images of actual faces, and the second based on simulated faces that allowed us to generate both expressions and their anti-expressions.

Experiment 1

For the first experiment, the images of emotional facial expressions were generated from the California facial expressions (CAFE) dataset (Dailey et al., 2001). The facial expressions used in this set had been certified according to the facial action coding system (FACS). Expressions of the six basic emotions and a neutral expression were selected from a single male individual in the CAFE dataset (individual 27, facial codes 027_n5, 027_a2, 027_d1, 027_f2, 027_h2, 027_m2, and 027_s1). These facial expressions were used for the neutral expression and to define the maximum intensity for each expression. The same individual was used for the adapt and test in order to maximize the strength of the expression aftereffects, which are selective for identity (Fox and Barton, 2007). While our results are thus restricted to a single identity, the highly stereotyped action patterns characterizing different expressions suggest that the pattern we observed is general.

All pictures were converted from the CAFE database into gray-scale bitmaps and presented at a size of 253×400 pixels. For each emotional expression, 101 graded intensities of the expression were created by morphing between the neutral facial expression and each basic expression using the Gryphon Software Corporation program MORPH Version 1.5 (see Figure 1). Sets of facial expressions were produced for each of the six basic emotions (anger, disgust, fear, happy, sad, and surprised), ranging in emotional intensity from 0 (the neutral face) to 100 (maximum intensity, corresponding to the original image of the expression).

Experiment 2

To create pairs of expression and anti-expressions, images of the emotional facial expressions were generated using the Singular

Inversions FaceGen Modeler program. This software is based on a 3D morphable model of faces, and details of the software and image set are described in O'Neil and Webster (2011). The program provides realistic portrayals of faces with varying identities and characteristics, including gender, ethnicity, age, and expression. Simulated faces from this program have been used in a number of other recent studies of face perception and adaptation (Shimojo et al., 2003; Russell et al., 2006; Schulte-Ruther et al., 2007; Oosterhof and Todorov, 2008; Potter and Corneille, 2008; O'Neil and Webster, 2011), and similar model faces have been found to convey information about expression that are reasonably comparable to images of actual faces (Dyck et al., 2008). One advantage of these modeled faces is that the strength of each of the basic expressions can be linearly titrated for a single fixed identity and pose, for which we chose a frontal view of an average Caucasian male face of 30 years as provided by FaceGen (Figure 2). A second advantage is that the strength can be varied in positive and negative directions to create both expressions and anti-expressions. That is, positive values produced the requested expression (e.g., anger) while negative values inverted the configural changes and thus yielded anti-expressions. For each pair, an array of 201 faces was created that ranged from the full anti-expression (intensity = -1) to the full original expression (intensity = +1).

PROCEDURE

For both experiments, stimuli were presented on a computer controlled CRT monitor. The face images subtended $\sim 7^\circ$ in height and were shown on a uniform background of $\sim 28^\circ$ by 37° . Subjects binocularly free-viewed the display from approximately 60 cm in an otherwise dark room, and responded using a hand-held keypad. They were asked to continuously view the adapting image but were not given specific instructions for viewing or fixation.

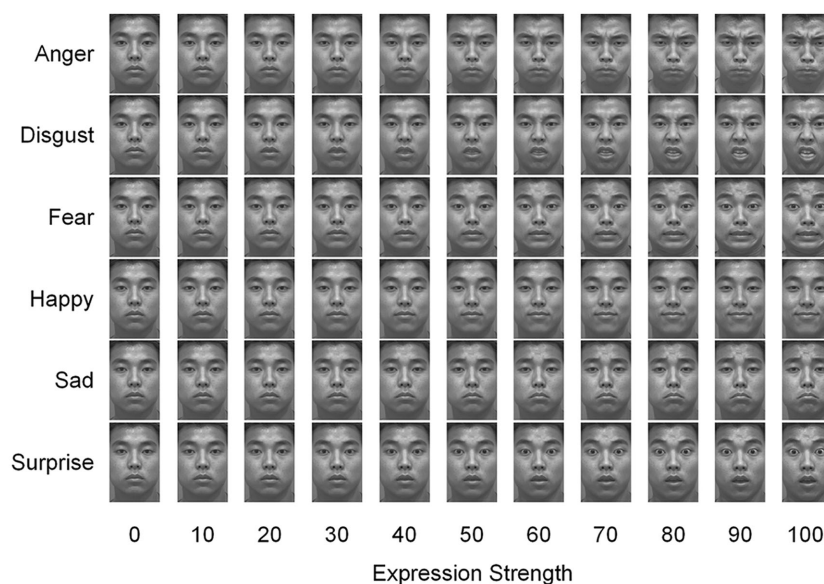


FIGURE 1 | A subset of the face image arrays used in Experiment 1. The face was varied from the neutral expression to one of the six basic expressions (corresponding to different rows). Successive images along each

row correspond to an increment of 10 units in expression strength formed by morphing between the neutral face (0) and the original posed expression (100).

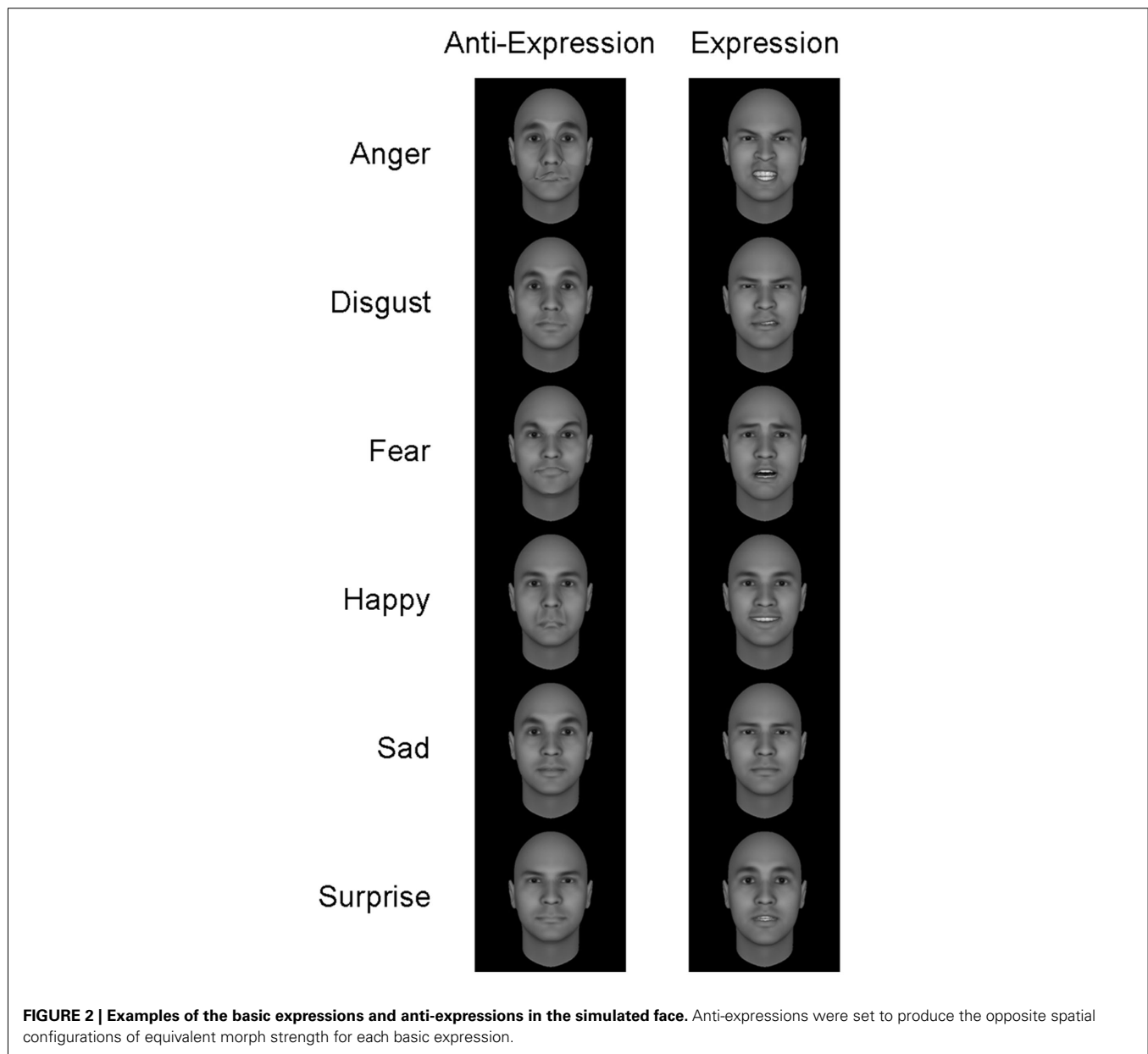


FIGURE 2 | Examples of the basic expressions and anti-expressions in the simulated face. Anti-expressions were set to produce the opposite spatial configurations of equivalent morph strength for each basic expression.

Experiment 1

The first experiment measured changes in the threshold intensity for recognizing different expressions after adapting to a given expression. In daily sessions lasting up to 1 h, observers were adapted to a single face image but were tested on all expressions, with the order of adapting expressions randomized across sessions. At the start of each run, the subject viewed the maximum intensity of one of the expressions or the neutral face for 5 min. Following this, a test face was presented for 1000 ms and then cycled with 3 s periods of readaptation, with the test and adapt images separated by 250 ms during which the screen was blank. The test face was drawn at random from one of the six expression categories, and the subject was thus required to make a six-alternative forced choice response to indicate the expression shown. The initial level along each category was chosen at random. Thresholds for identifying

each expression were found by varying subsequent levels with a staircase procedure. Six staircases were run simultaneously within each session, one for each expression, and settings continued until the staircase for each image set completed 10 reversals. Thresholds were estimated from the mean of the last seven reversals. In order to keep the task consistent throughout the run, when a staircase for a particular emotion terminated, the staircase continued, but the subject's responses for that staircase were no longer recorded.

Experiment 2

The second experiment tested for interactions between adaptation to each expression and its anti-expression. In a daily session subjects adapted to and made settings for only one expression. The task involved making a forced choice response to decide whether the presented face did or did not have the target expression. We

chose this over an alternative of asking which side of neutral the test face was on, since there were obvious asymmetries in the ability to classify positive or negative excursions. That is, while expressions were easy to identify, the anti-expressions were difficult to judge, precisely because they did not look like a basic expression (see **Figure 2**). Stimuli were varied in a staircase to estimate the category boundary based on the stimulus level at which the face was equally likely to be judged to have or not have the expression. Subjects made these settings after adapting to a neutral face or to the target expression or anti-expression shown at full strength, with the test face again varied by the staircase.

RESULTS

EXPRESSION RECOGNITION FOLLOWING ADAPTATION

Figure 3 plots for each expression the changes in recognition thresholds (i.e., the difference between the image levels required to correctly identify the presented expression after adapting to a given expression or to the neutral face). Positive values correspond to a higher threshold for the test expression and thus to loss in sensitivity to that expression, while negative values correspond to a reduced threshold and thus facilitation for the test expression. The results reveal that the aftereffects are strongly selective for the adapting expression. Specifically, the primary effect of the adaptation was to reduce recognition of the adapted expression, with little systematic effect on recognition for the unadapted expressions.

To evaluate these effects, the thresholds were compared with a two-way repeated measures ANOVA testing the variables of adapt category (seven levels including neutral) and test category (six levels). There was not a main effect of adapt expression [$F(6,66) = 1.23, p = 0.30$] but a significant effect for the test expression [$F(5,55) = 5.91, p < 0.001$]. Holm–Sidak comparisons revealed that this resulted because the absolute threshold for identifying fear in the face was higher than for disgust [$t(55) = 3.92, p = 0.0002$], happiness [$t(55) = 4.59, p < 0.0001$], or sadness [$t(55) = 4.53, p < 0.0001$]. (Note that these differences in absolute sensitivity to the expressions are not shown in **Figure 3**, which instead plots the *change* in the thresholds, i.e., the difference in thresholds after adapting to each expression vs. the neutral face). There were no other significant differences in absolute sensitivity to the different categories.

There was a significant interaction between the adapting and test categories [$F(30,330) = 10.05, p < 0.001$], and Holm–Sidak comparisons revealed strongly selective aftereffects for most expressions. Specifically, adaptation significantly altered the recognition thresholds only for the adapted expression for fear [$t(330) = 4.20, p < 0.001$], happiness [$t(330) = 3.86, p < 0.001$], sadness [$t(330) = 7.79, p < 0.001$], and surprise [$t(330) = 4.18, p < 0.001$]. Adapting to anger similarly increased the recognition threshold for anger [$t(330) = 5.52, p < 0.001$], though this also raised the threshold for sadness [$t(330) = 2.67, p = 0.008$]. The one exception to this pattern was thus for disgust, for which none of the aftereffects reached significance. Finally, all of the significant changes in the thresholds reflected a decrease in recognition after adaptation. That is, there was no case where adaptation to any expression enhanced the tendency to correctly identify an expression.

The results thus suggest that the aftereffects of adaptation to different facial expressions are highly selective for the adapting expression. In only one case was significant transfer observed (from adapting to anger on identifying sad). Moreover, the results failed to reveal any suggestion that adaptation to one expression facilitated the perception of a different expression; instead, almost all of the aftereffects are confined to a reduced response to the adapting axis. This suggests that – at least as probed by the present adaptation task – the representations of the different expressions are largely independent.

CHANGES IN PERCEIVED EXPRESSION FOLLOWING ADAPTATION TO EXPRESSIONS AND ANTI-EXPRESSIONS

As noted in the section “Materials and Methods,” in the second experiment subjects determined the stimulus level at which the target expression became visible after adapting to the neutral face or to either the expression or anti-expression. **Figure 4** plots for each category the changes in the settings (i.e., the setting when adapted to either the expression or anti-expression minus the setting when adapted to the neutral face). Large aftereffects are evident for most of the expressions. In particular, there is a clear trend for adaptation to each expression to make the target expression less visible, while adapting to the anti-expression induced the opposite change and thus made the expression more visible.

To evaluate these effects, the category boundaries were compared with a two-way repeated measures ANOVA testing the variables of adapt expression (six levels) and expression strength (three levels). There was a main effect of adapt expression [$F(5,30) = 6.66, p < 0.001$]. Holm–Sidak *a posteriori* comparisons revealed this resulted from the threshold for detecting a sad face being much larger than for the happy [$t(30) = 5.64, p < 0.001$]

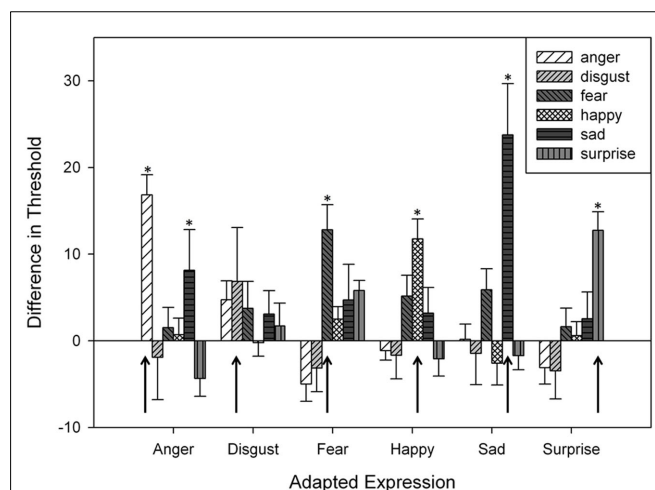


FIGURE 3 | Changes in recognition thresholds following adaptation (Experiment 1), averaged across 12 observers. Each bar plots the difference in the thresholds under adaptation to one of the basic expressions vs. for the neutral expression. Positive values correspond to a threshold increase. Each cluster of six bars corresponds to the six test expressions and a different adapting expression. Aftereffects when the adapt and test expression were the same are indicated by arrows. Asterisks indicate aftereffects that are significantly different from 0.

or angry expression [$t(30) = 3.60$, $p = 0.001$]. No other adapt expressions differed in their category boundaries.

There was also a main effect for expression strength [$F(2,12) = 71.71$, $p < 0.001$], with significant differences between all three conditions [all $t(12) > 3.36$, $p < 0.0057$]. There was no evidence of a significant adapt expression \times expression strength interaction [$F(10,60) = 1.40$, $p = 0.20$]. However, Holm–Sidak *a posteriori* comparisons revealed that the settings for the target expression differed from neutral for all expressions [all $t(60) > 4.74$, $p < 0.001$] while the anti-expression differed for all expressions [all $t(60) > 1.86$, $p < 0.035$] except for anger [$t(60) = 1.28$, $p = 0.11$] and fear [$t(60) = 1.57$, $p = 0.062$].

Finally, we also compared the size of the aftereffects for the expression and anti-expression faces with a two-way repeated measures ANOVA testing the variables of adapt expression (six levels) and expression sign (two levels, excluding neutral). There was a main effect of expression sign [$F(1,6) = 22.31$, $p = 0.003$], due to the aftereffects for the anti-expression being smaller than for the target expression. There was no evidence of a significant adapt expression \times expression strength interaction [$F(5,30) = 1.683$, $p = 0.169$]. However, Holm–Sidak *a posteriori* comparisons revealed that the aftereffects for the target expression were greater than the anti-expression for anger, disgust, fear, and surprise [all $t(30) > 2.409$, $p < 0.022$], while there was no observed difference in aftereffects for happy or sad [all $t(30) < 1.408$, $p > 0.169$].

Thus unlike the independence observed between adaptation to different actual expressions, each expression tended to show complementary aftereffects to the anti-expression. Thus the aftereffects for opposite facial configurations appeared yoked, in contrast to the different and at least conceptually complementary expressions of the basic emotions. However, for the conditions we tested

the anti-expression aftereffects were weaker than for the actual adapting expressions.

DISCUSSION

In this study we used adaptation to explore the visual representation of facial expressions. Consistent with previous work, we found that the perceived expression of a face can be strongly biased by prior adaptation to a facial expression (Russell and Fehr, 1987; Hsu and Young, 2004; Webster et al., 2004; Fox and Barton, 2007; Furl et al., 2007a,b; Benton and Burgess, 2008; Ellamil et al., 2008; Rutherford et al., 2008; Skinner and Benton, 2010; Cook et al., 2011; Pell and Richards, 2011). In our case these aftereffects were strongly selective for individual expressions. Specifically, adapting to an expression such as anger or happiness reduced sensitivity to anger or happiness in the face, while producing little change in sensitivity to other categories. Moreover, the changes in the thresholds for the adapting category did not lead to consistent increases in sensitivity to other categories. Thus the different basic expressions could be adapted largely independently. These results are consistent with the selective expression aftereffects reported by Skinner and Benton (2010) and Cook et al. (2011), and shows that this selectivity also occurs when the expressions and adaptation are probed relative to a neutral facial expression defined independently of the expression set. Again, the aftereffects relative to this neutral point are important for characterizing the selectivity of the adaptation, for the neutral expression may have a special status similar to the neutral identity that has been found to be important for defining the properties of face identity aftereffects (Rhodes and Jeffery, 2006).

Studies of face adaptation have varied widely in the strategies used to control for low-level or image-based aftereffects, for example between the local contours in the image. These steps include varying the size, position, or identity of the adapt and test stimuli (Webster and MacLeod, 2011). A limitation of our study was that we kept these parameters the same in order to maximize the strength of the adaptation, and thus the opportunities for interactions between the different expressions. While this could potentially have allowed the intrusion of lower-level aftereffects, these are strongly sensitive to spatial position (Xu et al., 2008), and have been found to be less evident when the faces are freely viewed without constraining fixation (Butler et al., 2008), as in our study. The nominally high-level aftereffects are themselves selective for position and size (Afraz and Cavanagh, 2008, 2009), and thus should also have been strongest when the adapt and test image were equated. Our stimuli should therefore have included a potential response change at higher levels where the image was represented as a face or expression. Thus adaptation-dependent interactions between different expressions arising at such sites should still have occurred, but were not observed in our conditions. On the other hand, it remains possible that aftereffects arising at early levels might mask a high-level aftereffect. Thus we cannot exclude the possibility that a different pattern of expression aftereffects might arise when the adapt and test faces share fewer image features. One argument against this is that our results again confirm the independence of different expression aftereffects reported by Skinner and Benton (2010), who included

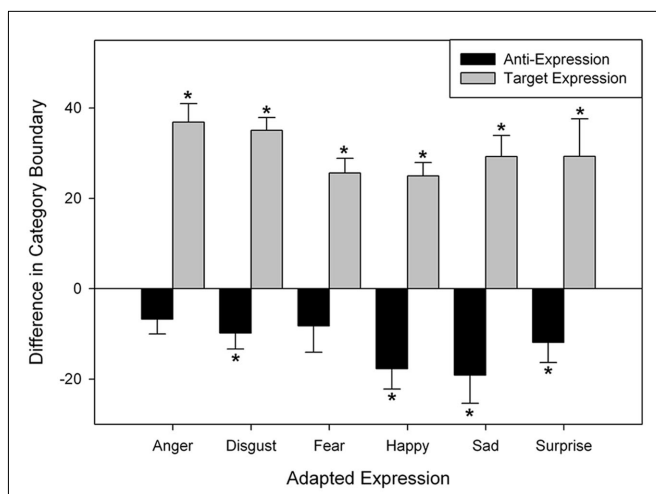


FIGURE 4 | Changes in the stimulus boundary for perceiving each basic expression after adapting to the expression or to the anti-expression (Experiment 2), based on the mean settings for seven observers. Each bar plots the difference between the stimulus level when adapted to the expressive vs. neutral face. Positive values indicate reduced sensitivity to the expression while negative values correspond to facilitation. Asterisks indicate aftereffects that are significantly different from 0.

a stationary fixation point but moving adapting image as a more explicit control for image-based adaptation.

While we observed little sign that adapting to one canonical expression facilitates an “opposite” expression, such interactions have been observed in previous studies. What could account for this difference? One case where interactions do clearly occur is when the stimuli are varied between two expressions, rather than in expression strength. For example, as noted in the Introduction, Webster et al. (2004) measured expression aftereffects in faces formed by morphing between two expressions such as happy and angry. Adapting to either expression caused the blended face to appear more like the unadapted expression. However, this composite face represented a mixture of two expressions rather than a neutral expression (which would only occur if two expressions were formed by opposite facial configurations). Thus their study probed the effects of adaptation on ambiguous expressions rather than neutral ones, and the fact that adaptation biased this ambiguity by selectively reducing sensitivity to one expression is consistent with the present findings. It is less certain how our results relate to the facilitation observed by Hsu and Young (2004), who measured sensitivity to expression in faces that varied between neutral and a given expression; or to Rutherford et al. (2008), who had subjects label the expressions perceived in a neutral face after adapting. In both cases prior adaptation to one expression made it more likely that the test faces would be labeled with a different expression. However, as we showed in Experiment 2, adaptation does in most cases alter the appearance of a neutral face – by inducing the opposite configural change in the face. This might cause a neutral face to appear more ambiguous, which could in turn increase the tendency to ascribe a different expression to it. Thus the facilitatory effects might not reflect a direct coupling between different categories. In any case, our results are similar to Hsu and Young (2004) in suggesting that any facilitation across expressions is substantially weaker than the reduction in sensitivity to the adapting expression, suggesting that any potential opponent-like couplings are correspondingly weaker.

A number of studies have examined the potential sites at which adaptation biases perceived expressions (Fox and Barton, 2007; Butler et al., 2008; Ellamil et al., 2008; Rutherford et al., 2008; Xu et al., 2008). As we noted in the Introduction, the aftereffects cannot be accounted for solely by low-level local features alone, such as the curvature of the mouth, or by high-level abstractions, such as emotional meaning. This suggests that the adaptation is acting partly at a site at which the expression is being represented in terms of its visual configuration. This “visual” locus may also explain why we observed little cross-talk between the different expression categories. The information for different expressions corresponds to combinations of changes in different facial features (Smith et al., 2005; Nusseck et al., 2008). Thus while different expressions might have conceptually opponent relationships (e.g., an individual is either happy or sad) the visual information conveying those states are not subject to the same constraints. Adaptation to the visual information in the face might therefore not reveal the functional relationships between the different emotional states conveyed by expression categories (Cook et al., 2011).

Our results are consistent with the possibility that this visual site of the adaptation may in part be prior to an explicit representation

of the expression, and thus occurs at a more generic level of the configural coding of the face. That is, at least part of the expression adaptation may act at a site common to many other facial attributes that have been examined with adaptation, by altering the representation of the spatial configuration of the face. In line with this, changing the facial configuration by distorting the image imbues the face with different expressions; (Ganel et al., 2004) and these distortions are highly adaptable (Webster and MacLin, 1999). Moreover, principal components analyses of facial variations point to distinct overlapping sources of variation between different identities or expressions, suggesting that at the visual level, expression, and generic shape are somewhat confounded (Calder and Young, 2005). It is also consistent with the finding that expression adaptation is selective for individual identity, so that whatever is adapted includes the shape information about identity, again arguing against a site where the expression has been explicitly extracted (Fox and Barton, 2007; Ellamil et al., 2008). (Intriguingly, the opposite has not been found. That is, identity adaptation completely transfers across a change in expression, suggesting that in this case adaptation might tap into a level where identity is coded independently of expression, Fox et al., 2008; or alternatively, it might conceivably act at a common level but information from this level is then pooled in different and asymmetric ways to form distinct representations of identity and expression).

In our study the primary evidence implicating a generic configural effect of the adaptation is from the aftereffects we found for anti-expressions. For most expressions, adapting to these faces also biased the appearance of the near-neutral face, yet these stimuli appear much more ambiguous and in this sense have less ecological validity than the basic expressions. If the adaptation were acting directly on processes coding expression then we might expect little response change from the anti-expressions, simply because these correspond to configural variations that are not clearly used to signal or detect expressions. However, as a change in facial shape they have a more equal status to an expression change, implying again that the adaptation may act at the level of the basic configural representation. In this regard our results again confirm the findings of Skinner and Benton (2010) in suggesting that the aftereffects reflect average shifts or a renormalization in the perceived expression of the face, consistent with a norm-based code of the type that has been suggested for invariant attributes of the face (Rhodes et al., 2005; Webster and MacLeod, 2011).

Are there also signs of an expression-specific site of the adaptation? One hint of this in our study was that the aftereffects for the anti-expressions were substantially weaker than for most of the basic expressions. This asymmetry is atypical of other reported aftereffects including facial distortions (Webster and MacLin, 1999; Rhodes et al., 2003; Watson and Clifford, 2003) and facial categories such as gender or ethnicity (Webster et al., 2004; Little et al., 2005; Ng et al., 2006; Jaquet et al., 2007; Jaquet and Rhodes, 2008) where opposites of the dimension appear to exert more equal effects on the neutral point. If mechanisms are sensitive only to the strength of a given expression – and if these mechanisms can be directly adapted – then the strongest response changes should occur only for faces with the appropriate expression.

However, there are a number of alternative accounts for the asymmetries we observed. First, because we measured when an

expression became apparent, the category boundary was always physically closer to the expression than the anti-expression. Thus the differences could in part reflect how far aftereffects to one level of the stimulus continuum spread to other levels – a local response shift would favor the locally closer expression. An argument against this is that the degree of asymmetry was not closely related to the threshold levels for detecting different expressions, and indeed was strongest for anger which had a relatively low threshold. Second, there were discrete qualitative changes on either side of the physically neutral face because most of the expressions include exposed teeth, while the complementary configurations did not. This could have provided a spatially local stimulus clue to the neutral point which might have been more impervious to adaptation, since it seems unlikely that a purely visual aftereffect to a closed mouth would affect the perception of the teeth. Given this difference it is surprising that strong aftereffects for anti-expressions were observed for some dimensions like happy faces which also included an open smile. Finally, the anti-expressions included changes in features such as eye brow thickness which could have introduced an apparent change in identity cues (though these changes corresponded to variations in the same identity with the brows raised or lowered). Moreover, for anger in particular, the full anti-expression included distortions which were outside the range of natural facial variations. We allowed this because these unnatural expressions nevertheless represented the equivalent opposing distortion in the linear model of the face, and because face aftereffects remain robust even when the adapting faces do not appear as plausible images of a real face (MacLin and Webster, 2001; Robbins et al., 2007; Seyama and Nagayama, 2009). Given these potential confounds, we cannot be certain of the basis for the asymmetries. Nevertheless, it is important to note that these stimulus asymmetries are inherent in the properties of actual facial expressions and not just in the stimuli we chose to probe them. That is – actual expressions often do include an open mouth that has no obvious facial counterpart in the anti-expression, and it is likely that there are not facial poses that are complementary and equal in intensity to the facial action patterns representing actual expressions. Thus the asymmetries are again at least consistent with sensitivity changes at expression-specific sites. And again, the facilitation found for most anti-expressions is inconsistent with changes only at these sites, and therefore also strongly implicates response changes at a more general level of configural coding.

We were motivated to explore the adaptation effects for facial expressions in part because there are only a small number of well-defined and salient dimensions to expressions. This differs from the perceptual attributes underlying facial identity, which remain very poorly defined in both number and form. This low-dimensional space offers the hope of quantifying the “tuning” properties for expression representations in the same way that adaptation has traditionally been used to characterize the channel selectivities of visual features such as color or form (Webster and MacLeod, 2011). What can our results say about these channels? On the one hand, in our case the different expressions do appear to be encoded largely independently. That is, to a first approximation the adapted level of the visual system appears to

represent the basic expressions as independent sources of information, and this is again consistent with the fact that as stimuli the basic expressions vary in independent ways (Smith et al., 2005; Nusseck et al., 2008) and also that we derive independent meanings from them. Our results thus support other evidence that different expressions are not encoded in terms of a common underlying framework (Calder et al., 2001). Yet on the other hand, our findings are not conclusive on whether the adaptation is producing sensitivity changes within mechanisms that are specifically tuned to the configurations defining different expressions. This is because we cannot exclude the possibility that the response changes are along an undefined set of dimensions which are in turn combined to form a representation of the expression (Cook et al., 2011). The latter is again hinted at by the fact that clear aftereffects occur for most of the anti-expression faces. These stimuli in fact present somewhat of a conundrum for modeling the adaptation (Webster and MacLeod, 2011). Similar to the types of models that have been developed to describe other facial aftereffects (Rhodes et al., 2005), representations of an expression might involve a balance between two mechanisms – one tuned to the expression and the other to the anti-expression. Yet the problem in this case is that the anti-expressions correspond to a set of stimuli that we rarely see, making it questionable that a mechanism would be built to detect them. And if the neutral face depends on how these two pools are balanced by adaptation, then the frequency differences mean that in the native state sensitivity should be strongly biased against the expression. Alternatively, a potential way out of this dilemma is again if the adaptation is acting at a more generic site coding different facial configurations. Processes for detecting a given expression could then be cobbled together from whatever dimensions might underlie the configural coding, without the need to build an opposing process. This leaves however the puzzling result that these opposing configurations generally lead to substantially weaker aftereffects. In any case, the point is that even for the simpler case of expressions it is not clear whether adaptation can be used to dissect the underlying channel structure.

Regardless of the possible sites of the response changes, adaptation may play important functional role in calibrating face perception, influencing judgments of expression as well as other attributes. One putative role of adaptation is to calibrate visual coding so that we can judge stimuli relative to a norm, and this role seems particularly relevant to expression perception since this involves detecting how the face deviates from neutral. A second possible role is to heighten sensitivity to these deviations by positioning the response to be maximally sensitive around the neutral point. Given that we are each exposed to different subsets of facial configurations that can be confounded with expressions (Neth and Martinez, 2008), adaptation may be critical for defining and maintaining an appropriate model of the neutral face. And given the social importance of facial expressions, this adaptation would also be critical for our ability to derive meaning from the face.

ACKNOWLEDGMENTS

Supported by National Eye Institute Grant EY-10834.

REFERENCES

- Adolphs, R., Damasio, H., Tranel, D., and Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *J. Neurosci.* 16, 7678–7687.
- Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372, 669–672.
- Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. R. (1995). Fear and the human amygdala. *J. Neurosci.* 15, 5879–5891.
- Afraz, A., and Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *J. Vis.* 9, 11–17.
- Afraz, S. R., and Cavanagh, P. (2008). Retinotopy of the face aftereffect. *Vision Res.* 48, 42–54.
- Benton, C., and Burgess, E. (2008). The direction of measured face aftereffects. *J. Vis.* 8, 1–6.
- Butler, A., Oruc, I., Fox, C. J., and Barton, J. J. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Res.* 1191, 116–126.
- Calder, A., and Young, A. (2005). Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651.
- Calder, A. J., Lawrence, A. D., and Young, A. W. (2001). Neuropsychology of fear and loathing. *Nat. Rev. Neurosci.* 2, 352–363.
- Cook, R., Matei, M., and Johnston, A. (2011). Exploring expression space: adaptation to orthogonal and anti-expressions. *J. Vis.* 11(4): 2, 1–9.
- Dailey, M. N., Cottrell, G. W., and Reilly, J. (2001). *California Facial Expressions (CAFE)*. La Jolla, CA: Computer Science and Engineering Department, University of California San Diego.
- Dyck, M., Winbeck, M., Leiber, S., Chen, Y., Gur, R. C., and Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLoS ONE* 3, e3628. doi:10.1371/journal.pone.0003628
- Ekman, P. (1992). Are there basic emotions? *Psychol. Rev.* 99, 550–553.
- Ekman, P., and Friesen, W. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ellamil, M., Susskind, J. M., and Anderson, A. K. (2008). Examinations of identity invariance in facial expression adaptation. *Cogn. Affect. Behav. Neurosci.* 8, 273–281.
- Fox, C. J., and Barton, J. J. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Res.* 1127, 80–89.
- Fox, C. J., Oruc, I., and Barton, J. J. S. (2008). It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *J. Vis.* 8(3): 11, 1–13.
- Furl, N., van Rijsbergen, N. J., Treves, A., and Dolan, R. J. (2007a). Face adaptation aftereffects reveal anterior medial temporal cortex role in high level category representation. *Neuroimage* 37, 300–310.
- Furl, N., van Rijsbergen, N. J., Treves, A., Friston, K. J., and Dolan, R. J. (2007b). Experience-dependent coding of facial expression in superior temporal sulcus. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13485–13489.
- Ganel, T., Goshen-Gottstein, Y., and Ganel, T. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: the parallel-route hypothesis revisited. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 583–597.
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci. (Regul. Ed.)* 4, 223–233.
- Hsu, S. M., and Young, A. W. (2004). Adaptation effects in facial expression recognition. *Vis. Cogn.* 11, 871–899.
- Jaquet, E., and Rhodes, G. (2008). Face aftereffects indicate dissociable, but not distinct, coding of male and female faces. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 101–112.
- Jaquet, E., Rhodes, G., and Hayward, W. G. (2007). Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences. *Q. J. Exp. Psychol.* 60, 1457–1467.
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302.
- Kesler/West, M. L., Andersen, A. H., Smith, C. D., Avison, M. J., Davis, C. E., Kryscio, R. J., and Blonder, L. X. (2001). Neural substrates of facial emotion processing using fMRI. *Brain Res. Cogn. Brain Res.* 11, 213–226.
- Leopold, D. A., O'Toole, A. J., Vetter, T., and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat. Neurosci.* 4, 89–94.
- Little, A. C., DeBruine, L. M., and Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proc. Biol. Sci.* 272, 2283–2287.
- MacLin, O. H., and Webster, M. A. (2001). Influence of adaptation on the perception of distortions in natural images. *J. Electron. Imaging* 10, 100–109.
- Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D., Young, A. W., Calder, A. J., and Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain* 121, 47–57.
- Neth, D., and Martinez, A. (2008). Emotion perception in emotionless face images suggests a norm-based representation. *J. Vis.* 9, 1–11.
- Ng, M., Ciaramitaro, V. M., Anstis, S., Boynton, G. M., and 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.
- Nusseck, M., Cunningham, D. W., Wallraven, C., and Bulthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *J. Vis.* 8, 1–23.
- O'Neil, S., and Webster, M. A. (2011). Adaptation and the perception of facial age. *Vis. Cogn.* 19, 534–550.
- Oosterhof, N. N., and Todorov, A. (2008). The functional basis of face evaluation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11087–11092.
- Pell, P. J., and Richards, A. (2011). Cross-emotion facial expression aftereffects. *Vision Res.* 51, 1889–1896.
- Plutchik, R. (2001). The nature of emotions. *Am. Sci.* 89, 344–350.
- Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* 17, 715–734.
- Potter, T., and Corneille, O. (2008). Locating attractiveness in the face space: faces are more attractive when closer to their group prototype. *Psychon. Bull. Rev.* 15, 615–622.
- Rhodes, G., and Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Res.* 46, 2977–2987.
- Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., and Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychol. Sci.* 14, 558–566.
- Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., and Clifford, C. W. G. (2005). “Adaptation and face perception – how aftereffects implicate norm based coding of faces,” in *Fitting the Mind to the World: Adaptation and Aftereffects in High-Level Vision*, eds C. W. G. Clifford and G. Rhodes (Oxford: Oxford University Press), 213–240.
- Robbins, R., McKone, E., and Edwards, M. (2007). Aftereffects for face attributes with different natural variability: adapter position effects and neural models. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 570–592.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., and Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126(Pt 11), 2381–2395.
- Russell, J., and Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *J. Exp. Psychol. Gen.* 116, 223–237.
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Social Psychol.* 39, 1161–1178.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychol. Bull.* 115, 102–141.
- Russell, R., Sinha, P., Biederman, I., and Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception* 35, 749–759.
- Rutherford, M. D., Chattha, H. M., and Krysko, K. M. (2008). The use of aftereffects in the study of relationships among emotion categories. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 27–40.
- Said, C. P., Haxby, J. V., and Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1660–1670.
- Schulte-Ruther, M., Markowitsch, H. J., Fink, G. R., and Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. *J. Cogn. Neurosci.* 19, 1354–1372.
- Seyama, J., and Nagayama, R. S. (2009). Probing the uncanny valley with the eye size aftereffect. *Presence (Camb.)* 18, 321–339.
- Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 1317–1322.

- Skinner, A. L., and Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychol. Sci.* 21, 1248–1253.
- Smith, M. L., Cottrell, G. W., Gosselin, F., and Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychol. Sci.* 16, 184–189.
- Sprengelmeyer, R., Rausch, M., Eysel, U. T., and Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. *Proc. R. Soc. Lond. B Biol. Sci.* 26, 1927–1931.
- Watson, T. L., and Clifford, C. W. G. (2003). Pulling faces: an investigation of the face-distortion aftereffect. *Perception* 32, 1109–1116.
- Webster, M. A. (2011). Adaptation and visual coding. *J. Vis.* 11(5): 3, 1–23.
- Webster, M. A., Kaping, D., Mizokami, Y., and Duhamel, P. (2004). Adaptation to natural facial categories. *Nature* 428, 558–561.
- Webster, M. A., and MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1702–1725.
- Webster, M. A., and MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychon. Bull. Rev.* 6, 647–653.
- Xu, H., Dayan, P., Lipkin, R. M., and Qian, N. (2008). Adaptation across the cortical hierarchy: low-level curve adaptation affects high-level facial-expression judgments. *J. Neurosci.* 28, 3374–3383.

04 January 2012; published online: 24 January 2012.

Citation: Juricevic I and Webster MA (2012) Selectivity of face aftereffects for expressions and anti-expressions. *Front. Psychology* 3:4. doi: 10.3389/fpsyg.2012.00004

This article was submitted to *Frontiers in Perception Science*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Juricevic and Webster. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 October 2011; paper pending published: 06 December 2011; accepted: