

# How is facial expression coded?

**Nichola Burton**

ARC Centre of Excellence in Cognition and its Disorders,  
School of Psychology, University of Western Australia,  
Crawley, WA, Australia



**Linda Jeffery**

ARC Centre of Excellence in Cognition and its Disorders,  
School of Psychology, University of Western Australia,  
Crawley, WA, Australia



**Andrew J. Calder**

ARC Centre of Excellence in Cognition and its Disorders,  
School of Psychology, University of Western Australia,  
Crawley, WA, Australia  
MRC Cognition and Brain Sciences Unit, Cambridge, UK



**Gillian Rhodes**

ARC Centre of Excellence in Cognition and its Disorders,  
School of Psychology, University of Western Australia,  
Crawley, WA, Australia



Facial expression is theorized to be visually represented in a multidimensional expression space, relative to a norm. This norm-based coding is typically argued to be implemented by a two-pool opponent coding system. However, the evidence supporting the opponent coding of expression cannot rule out the presence of a third channel tuned to the center of each coded dimension. Here we used a paradigm not previously applied to facial expression to determine whether a central-channel model is necessary to explain expression coding. Participants identified expressions taken from a fear/antifear trajectory, first at baseline and then in two adaptation conditions. In one condition, participants adapted to the expression at the center of the trajectory. In the other condition, participants adapted to alternating images from the two ends of the trajectory. The range of expressions that participants perceived as lying at the center of the trajectory narrowed in both conditions, a pattern that is not predicted by the opponent-coding model but can be explained by the central-channel model. Adaptation to the center of the trajectory also increased identification of both fear and antifear, which may indicate a functional benefit for adaptive coding of facial expression.

## Introduction

We are able to extract many types of information from a face, including the person's identity, gender, race, expression, and numerous other attributes. This ability requires sensitivity to small differences between faces, impressive given the similarity of faces as visual patterns. For instance, we are able to use subtle changes in the arrangement of the features of a face to perceive a person's expression, which is often a source of important social cues. This sensitivity has led to great interest in the visual coding mechanisms that underlie the representation of facial expression and other types of facial information.

Many aspects of face perception (e.g., expression, identity, race, gender) are theorized to be visually coded relative to a norm in face space. Each dimension in face space represents a way in which faces are perceived to vary (although the dimensions used are not yet known). The norm represents the central tendency of previously seen faces and so lies at the center of this space. The positions of new faces in the space are coded relative to this norm. In this way, the coding of faces captures what is different or distinctive about them, which may contribute to our excellent face perception ability. Norm-based coding is also adaptive; the norm is constantly updated to represent the range of faces we

Citation: Burton, N., Jeffery, L., Calder, A. J., & Rhodes, G. (2015). How is facial expression coded? *Journal of Vision*, 15(1):1, 1–13, <http://www.journalofvision.org/content/15/1/1>, doi:10.1167/15.1.1.

encounter. This adaptability may help to calibrate the system to the range of faces that are most prevalent, optimizing sensitivity across that range (for a review, see Rhodes & Leopold, 2011; Webster & MacLeod, 2011).

Evidence in support of norm-based coding of facial expression comes from paradigms using adaptation aftereffects (Burton, Jeffery, Skinner, Benton, & Rhodes, 2013; Cook, Matei, & Johnston, 2011; Skinner & Benton, 2010, 2012). An aftereffect occurs when viewing a stimulus alters participants' perception of subsequent stimuli. The responses of neural populations that were initially stimulated by the stimulus begin to be suppressed with exposure. This reduction in responsiveness relative to other neural pools causes the percepts of subsequently viewed stimuli to be biased away from the adaptor. The greater the initial response, the larger the subsequent suppression and so the greater the aftereffect (Maddess, McCourt, Blakeslee, & Cunningham, 1988; Movshon & Lennie, 1979). Examining the size and direction of the aftereffects produced by adapting to particular expressions allows us to test hypotheses about the neural populations that code expression.

Norm-based face coding is often theorized to be instantiated by an opponent coding system (e.g., Rhodes & Jeffery, 2006; Rhodes et al., 2005; Robbins, McKone, & Edwards, 2007; Tsao & Freiwald, 2006). In this type of system, there are two pools of neurons that code a given perceptual dimension: one pool that responds maximally to one extreme of the dimension and one pool that responds maximally to the other extreme. The norm is implicitly coded as the point at which the two pools respond equally. This model allows for efficient coding; the maximal response is reserved for aspects of a face that are particularly distinctive and so useful for recognition with less energy devoted to representing the less useful aspects that are common to most faces. In the opponent-coding model, adaptation is theorized to shift the position of the norm by altering the point at which both neural pools respond equally. It is this shift in the norm that creates the aftereffect.

Evidence for the opponent coding of facial expression comes chiefly from the near–far aftereffect paradigm (Burton et al., 2013; Skinner & Benton, 2010, 2012). In the opponent-coding model, adaptors that lie further from the norm will produce larger aftereffects than adaptors that lie closer to the norm (Robbins et al., 2007). This occurs because the more extreme adaptors result in stronger neural suppression and a larger shift in the norm than less extreme adaptors.

The adaptors used in the near–far paradigm are antiexpressions, produced by morphing a face along a trajectory that runs from an expression through the average expression (a morph-average of the basic

expressions taken as an approximation of the norm, which is the central tendency of expressions that the participant has previously seen) and beyond the average to a point of equal distance beyond it. This antiexpression differs from the average to the same extent as the original expression, but in the opposite direction (so raised eyebrows become lowered, for instance). This expression/antiexpression trajectory does not necessarily correspond to an underlying coding dimension in face space. However, perception of faces along this trajectory will activate underlying expression-relevant dimensions, so we can use adaptation on the trajectory to examine the coding of those dimensions. Adapting to an antiexpression biases perception toward the original expression; for instance, adapting to antifear produces an aftereffect that biases perception toward fear (Burton et al., 2013; Skinner & Benton, 2010, 2012). As predicted by the opponent-coding model, more extreme antiexpression adaptors create stronger aftereffects than less extreme antiexpression adaptors (Burton et al., 2013; Skinner & Benton, 2010, 2012).

The near–far paradigm may not, however, be sufficient to rule out alternative models with more than two channels. Narrowband multichannel models can produce an initial increase in aftereffects followed by a subsequent decline (Blakemore & Sutton, 1969; Clifford, Wenderoth, & Spehar, 2000). A central-channel model with widely tuned channels coding the ends of the dimension but with an additional channel coding the center of the dimension (as described by Calder, Jenkins, Cassel, & Clifford, 2008; Lawson, Clifford, & Calder, 2009, 2011) would also predict an initial increase in aftereffects. Therefore, we cannot currently rule out alternative, nonopponent models of expression coding.

There is evidence that a nonopponent, multichannel model with a central channel rather than opponent coding is used to code gaze direction (Calder et al., 2008) and head and body orientation (Lawson et al., 2009, 2011). Those studies used a different paradigm, in which participants classified stimuli taken from along a trajectory using three labels: one for each end of the trajectory and one for the center of the trajectory (e.g., “leftward gaze,” “rightward gaze,” and “direct gaze” in the case of a gaze trajectory). Baseline judgments (i.e., without adaptation) are compared to judgments made in two adaptation conditions: one in which participants adapt to the center of the trajectory and one in which participants adapt to alternating images of the ends of the trajectory. The dependent variable of interest is the range of stimuli that are judged to belong to the central category: the “central range.” If we assume that adaptation results in a suppression of neural pools stimulated by the adaptor, then we can derive predictions about the effect of adaptation on the

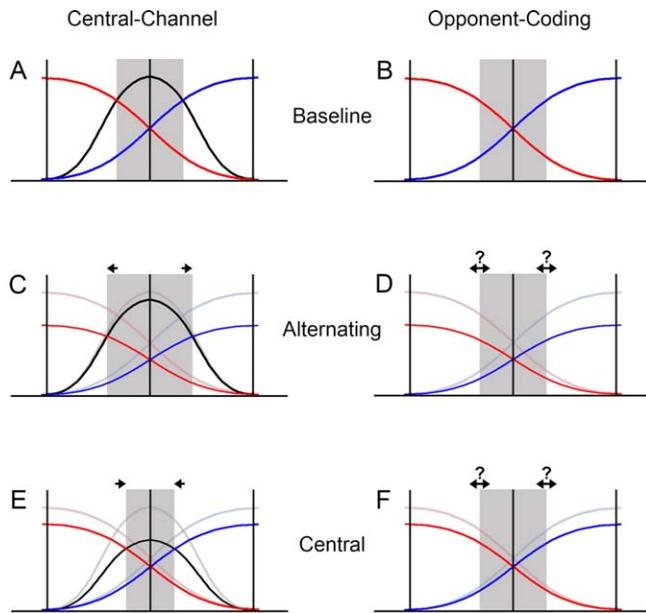


Figure 1. In a central-channel model (A), adapting to both endpoints of the dimension will increase the range of faces seen as central, shown here in gray (C); adapting to the center of the dimension will reduce the range of faces seen as central (E). In an opponent-coded model (B), adapting to both endpoints (D) will shift the central range in the same direction as adapting to the center (F). Figure adapted with permission from Lawson et al. (2011).

central range from an opponent-coding model that differ from the predictions derived from a central-channel model.

In a central-channel model, alternating adaptation should narrow the central range because the outer pools become relatively less responsive (Figure 1C). In contrast, central adaptation should widen the central range because the central channel becomes relatively less responsive (Figure 1E). This pattern was found for gaze direction and for head and body orientation. For instance, adapting to a central (front-facing) head direction reduced the range of directions perceived as front-facing (narrowed the central range), and adapting to alternating left and right head directions increased the range of directions perceived as front-facing (widened the central range) (Lawson et al., 2011).

The opponent-coding model does not predict these opposing changes in the size of the central range. It is difficult to predict exactly what will happen to the size of the central range following adaptation in the case of opponent coding. Whether adaptation results in a narrowing or widening of the central range depends on the criterion by which a stimulus is perceived as central and the shape of the response curves of the two pools (see Lawson et al., 2011, for further discussion). However, any change in the size of the central range should be in the same direction for both adaptation

conditions. Alternating adaptation stimulates the two opponent pools, and central adaptation also stimulates these two opponent pools although possibly to a lesser extent. For this reason, we can expect that any effect of adaptation seen in the alternating condition (whether widening or narrowing the central range) should also be seen in the central condition (Figure 1D, F). Thus, in an opponent-coding model, we do not expect the opposing changes in the size of the central range predicted by the central-channel model.

We used this paradigm to determine which of the two models better describes the coding of facial expressions. We showed participants expressions from a morphed expression trajectory that ran from a fear expression through the average expression and out to an antifear expression. We chose this trajectory because fear and antifear are distinctive and would be easy for participants to learn. We taught participants to use arbitrary labels to identify the expressions at each end of the trajectory (“A” for antifear and “C” for fear) and the center of the trajectory (“B”). Participants used these labels to classify faces taken from along the expression trajectory at baseline (no adaptation) and in two adaptation conditions: central adaptation, in which participants adapted to “B,” and alternating adaptation, in which participants adapted to alternating images of “A” and “C.” Our dependent variable was the central range, the range of levels labeled “B.” Widening of the central range after alternating adaptation and narrowing of the central range after central adaptation (relative to baseline) would support a central-channel model. This opposing change cannot be explained by an opponent-coding model, which predicts that any changes in the central range would occur in the same direction for both adaptation conditions.

Our experimental paradigm also allowed us to address an additional question: whether there is a functional benefit of adaptation for expression perception. In low-level vision, adaptive coding helps to calibrate the limited resources of the coding system to the current range of stimuli (Clifford, 2002; Clifford & Rhodes, 2005; Thompson & Burr, 2009), but evidence of this benefit in face perception has been mixed (for reviews, see Armann, Jeffery, Calder, Bülhoff, & Rhodes, 2011; Rhodes & Leopold, 2011). So far, no research has examined a possible functional benefit of adaptive coding of facial expression.

Both the central-channel and opponent-coding models can accommodate an improvement in participants’ identification of expressions following adaptation. In the central-channel model, increased identification of fear and antifear following central adaptation can be explained by the suppression of the central neural channel (Figure 1E). In the opponent-coding model, identification of fear and antifear may be



Figure 2. The expressions defining the test trajectory from left to right:  $-100\%$  (antifear),  $0\%$  (average), and  $100\%$  (fear). For ease of participant response, these expressions were labeled A, B, and C, respectively.

increased if adaptation steepens the tuning functions of the two pools, altering their relative responses. Thus, regardless of which model is supported by our results, we may find evidence of a functional benefit of expression adaptation.

## Method

### Participants

Twenty-four Caucasian participants were recruited from the University of Western Australia. This sample size was judged to be sufficient based on the size of samples used in research utilizing this paradigm with other stimuli (Calder et al., 2008; Lawson et al., 2009, 2011) and in research that has found significant aftereffects using expression stimuli (Burton et al., 2013; Skinner & Benton, 2010, 2012). One participant's data were excluded from analysis due to a computer error during testing. The remaining 23 participants (five male) had a mean age of 20.7 years,  $SD = 5.4$  years. Participants were either awarded credit as part of a psychology course or were reimbursed \$15 for travel expenses.

### Stimuli

Stimuli were adapted from those used by Skinner and Benton (2010, 2012). These were gender-neutral expressive faces created from images of 20 Caucasian individuals (10 male and 10 female) posing various expressions. For each expression, an average was created from the 20 images of that expression using morphing software (see Skinner & Benton, 2010, for more details). An overall average expression was created by taking an average of seven of these gender-neutral expressions (happy, sad, angry, fearful, surprised, disgusted, and neutral).

The gender-neutral fear expression lies at one end of the test trajectory ( $100\%$ ); the other end of the trajectory was created by morphing through the average expression ( $0\%$ ) and out into antifear ( $-100\%$ ), which differs from the average to the same extent as fear but in the opposite direction (see Figure 2). The  $100\%$ ,  $0\%$ , and  $-100\%$  expressions were our target expressions, also used as adaptors. The three expressions were labeled A ( $-100\%$ ), B ( $0\%$ ), and C ( $100\%$ ) for ease of participant response. The test stimuli were faces from nine points along the trajectory:  $-80\%$ ,  $-60\%$ ,  $-40\%$ ,  $-20\%$ ,  $0\%$ ,  $20\%$ ,  $40\%$ ,  $60\%$ , and  $80\%$  (Figure 3). To reduce testing duration, the  $-100\%$  and  $100\%$  expressions were not included as test faces.

Stimuli were shown in gray scale on a 21.5-in. iMac monitor at a viewing distance of approximately 50 cm. Adaptors subtended a visual angle of  $8.9^\circ \times 12.1^\circ$ . Test stimuli were shown at  $75\%$  of that size ( $6.9^\circ \times 9.1^\circ$ ) to reduce the contribution of retinotopic adaptation.

### Procedure

Participants began each testing session with a training task that taught them to identify the three target expressions by their labels (A, B, C). Participants were first introduced to the expressions and their associated labels on-screen. They were then shown an expression on-screen and were given unlimited time to identify it using a marked keyboard key. A brief tone indicated whether the response was correct or not, and the next expression was shown. When participants were able to correctly identify a random sequence of nine expressions (each target expression appearing three times), they moved to the next training phase.

In the next training phase, participants saw the expressions for only 200 ms. Each expression was followed by a 150-ms blank interstimulus interval (ISI). Participants then saw a response screen (“?”) and responded as before. In this phase, there was no feedback. Participants were again required to correctly identify a sequence of nine expressions to move on. If they went through three incorrect repetitions of this sequence, they returned to the previous training phase and worked through that again before coming back to the no-feedback training phase. Nine participants were required to return to the feedback phase in the first session, and eight participants were required to return to the feedback phase in the second session. Participants completed a mean of 2.7 total repetitions of the no-feedback training phase in the first session ( $SD = 2.1$ ) and a mean of 2.3 total repetitions of the no-feedback training phase in the second session ( $SD = 1.7$ ).<sup>1</sup>

The main testing procedure was adapted from Lawson et al. (2009). There were five testing phases:

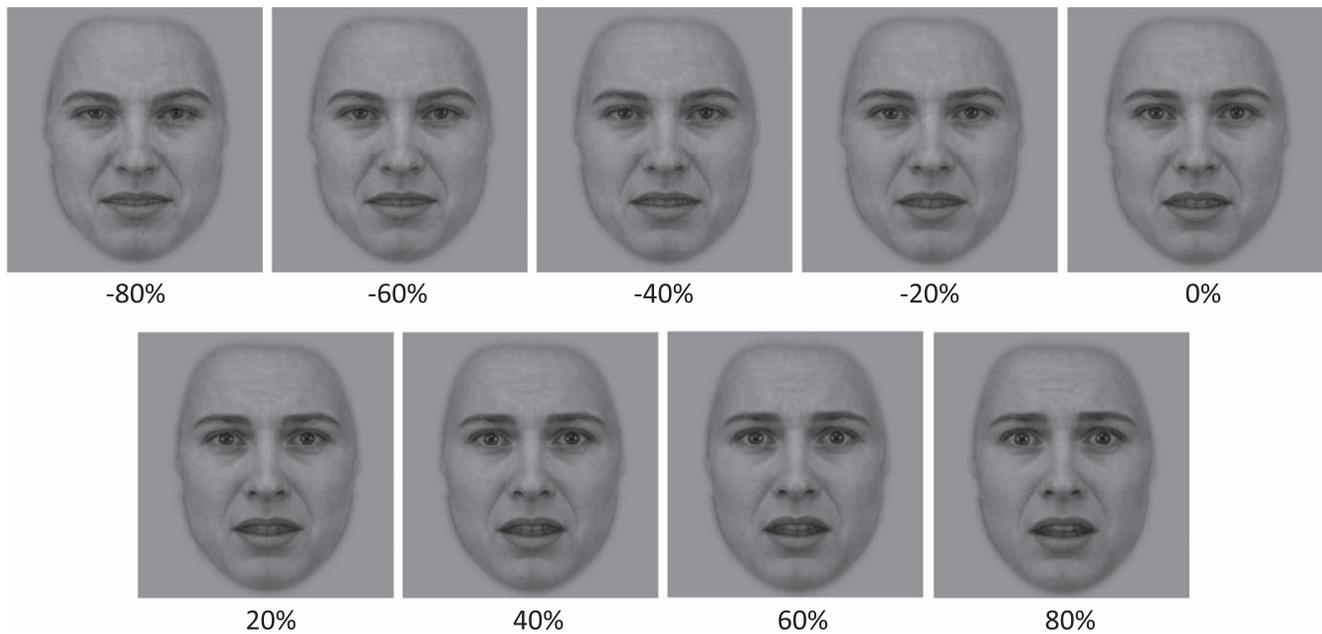


Figure 3. The nine test expression levels, ranging from antifear (−80%) through the average expression (0%) to fear (80%).

baseline, first adaptation, baseline, second adaptation, baseline. In the baseline phase, participants were shown a test face for 200 ms, followed by a 150-ms ISI. Participants then identified which target expression they had seen using the labeled keys. Participants completed six of these trials for each of the nine test levels in random order (54 trials total). In the alternating adaptation phase, participants first adapted to 40 alternating images of the −100% and 100% expressions (20 of each), each shown for 4000 ms and separated by 200-ms blank ISIs (total adaptation time of 160 s). They then completed 54 trials of identifying test faces as in the baseline but with each test face preceded by six alternating top-up adaptor images, each shown for 1000 ms and separated by 200-ms ISIs. The central adaptation condition followed the same procedure, but instead of the initial alternating adaptors, participants adapted to the 0% expression for 40 repeated 4000-ms exposures separated by 200-ms blank ISIs (total adaptation time of 160 s), and the top-up images were six 1000-ms exposures of the 0%



Figure 4. The brightness changes used in the attention task from left to right: the original 0% expression, with the eyes brightened, with the lips brightened.

expression separated by 200-ms ISIs. Participants were allowed to move their eyes freely throughout the task.

To maintain attention during the long initial adaptation phase, we included a secondary attention task. Over the course of some of the 4000-ms-long adaptor exposures either the irises or lips would become brighter (see Figure 4). This brightness change occurred in steps over the last 1000 ms of the exposure: Five images were shown with the feature increasingly brightened; the first four brightness levels were shown for 125 ms each with the final, brightest image left on-screen for the remaining 500 ms. Before adaptation began, participants were shown examples of the brightened eyes and lips and were then shown a sequence of faces (4000-ms exposures with 150-ms ISIs) in which either the eyes or the lips changed as described above. Participants pressed a marked key as soon as they saw a change. Participants practiced this task until they and the experimenter were comfortable that they understood what was required. Of the 40 adaptation exposures, eight contained an eye change and eight contained a lip change. Participants indicated whether the eyes or the lips changed using marked keys as soon as they saw a change. In the alternating adaptation phase, changes were equally distributed across the −100% and 100% adaptors.

Participants began with two blocks of baseline testing (54 trials in each). They then learned to identify the eye and lip changes for the attention task. Next participants completed one block of adaptation testing (either alternating or central). This was followed by two more blocks of baseline testing. Participants were reintroduced to the eye and lip changes and completed

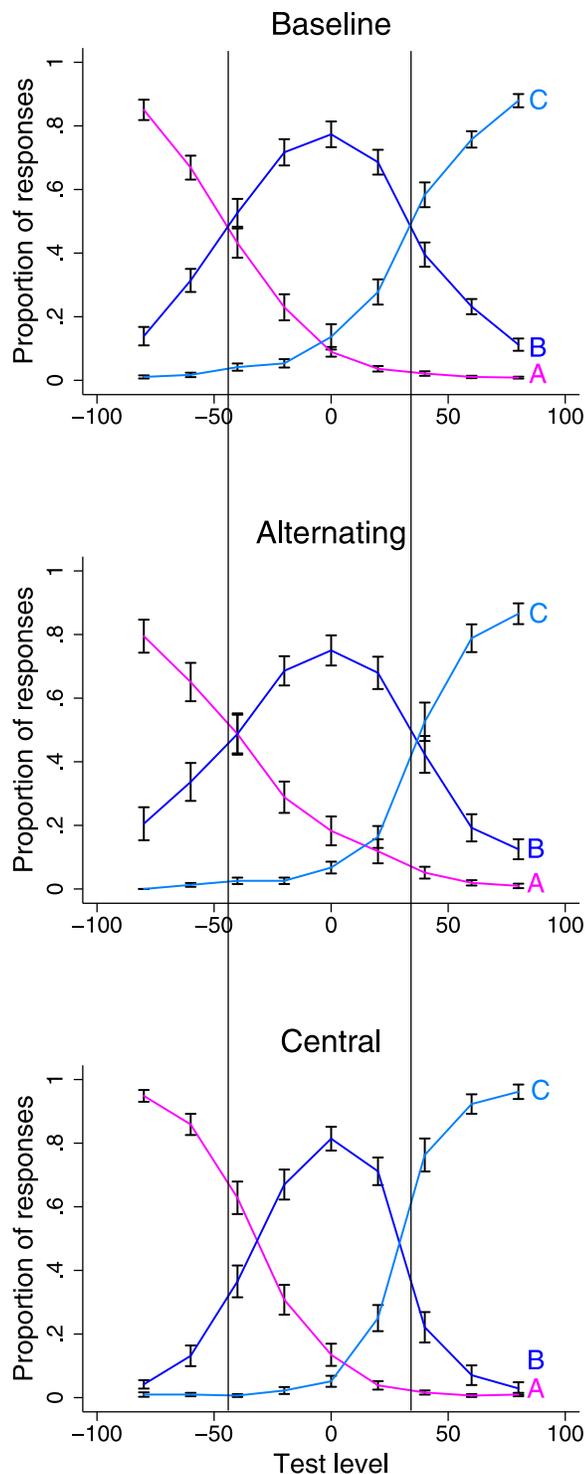


Figure 5. Mean proportion of “A” (antifear, shown in magenta), “B” (average, shown in dark blue), and “C” (fear, shown in light blue) responses to each level of the test trajectory in the baseline, alternating adaptation, and central adaptation conditions across all participants. To aid comparison across adaptation conditions, we show vertical lines that indicate the test level of the A-B and B-C crossing points in the baseline condition.

a second block of adaptation testing (whichever form they had not previously completed). Finally, there were two more blocks of baseline testing.

Participants completed two of these testing sessions, each beginning with the training task. Participants either completed the alternating adaptation phase or the central adaptation phase first in both of their sessions; the order was counterbalanced between participants. Each session took approximately 45 min to complete. The two sessions were completed between 1 and 28 days apart ( $M = 6.48$  days,  $SD = 5.35$  days).

## Results

Baseline data were collected in pairs of blocks before, between, and after the two adaptation blocks. Following Lawson et al. (2009), we discarded the data from the first baseline block of each pair: from the first pair to allow a practice period and from subsequent pairs to allow any lingering adaptation to dissipate as much as possible. The responses in the remaining baseline blocks (two, four, and six) were taken as a measure of performance in the absence of adaptation. The two sessions of testing were collapsed together for analysis.

We aimed to determine whether a central-channel model is necessary for the coding of facial expression. To do this, we compared the range of expressions that participants classified as central (“B”) at baseline to the range classified as central in each of our two adaptation conditions. The central-channel model predicts that this range should widen relative to baseline after alternating adaptation and narrow relative to baseline after central adaptation. The opponent-coding model predicts that any changes in the range should be in the same direction for both adaptation conditions.

Mean proportions of A, B, and C responses at each test strength in each adaptation condition are given in Figure 5. Examples of individual data from two participants are given in Figure 6. Visual inspection of the graphs indicates that the range of expression over which participants were more likely to respond “B” than “A” or “C” became narrower after adaptation in both conditions relative to baseline.

In order to statistically examine the effect of adaptation on the size of the central range, we fit a mixed multinomial logit model to participants’ responses (see Supplementary Materials for details). This model uses the adaptation condition (baseline, alternating, central) to estimate the probability of selecting a given label (“A”/“B”/“C”) in response to a given test level. The mixed logit model is an extension of standard logistic regression but allows the parameters of the fit functions to vary across participants. This feature of

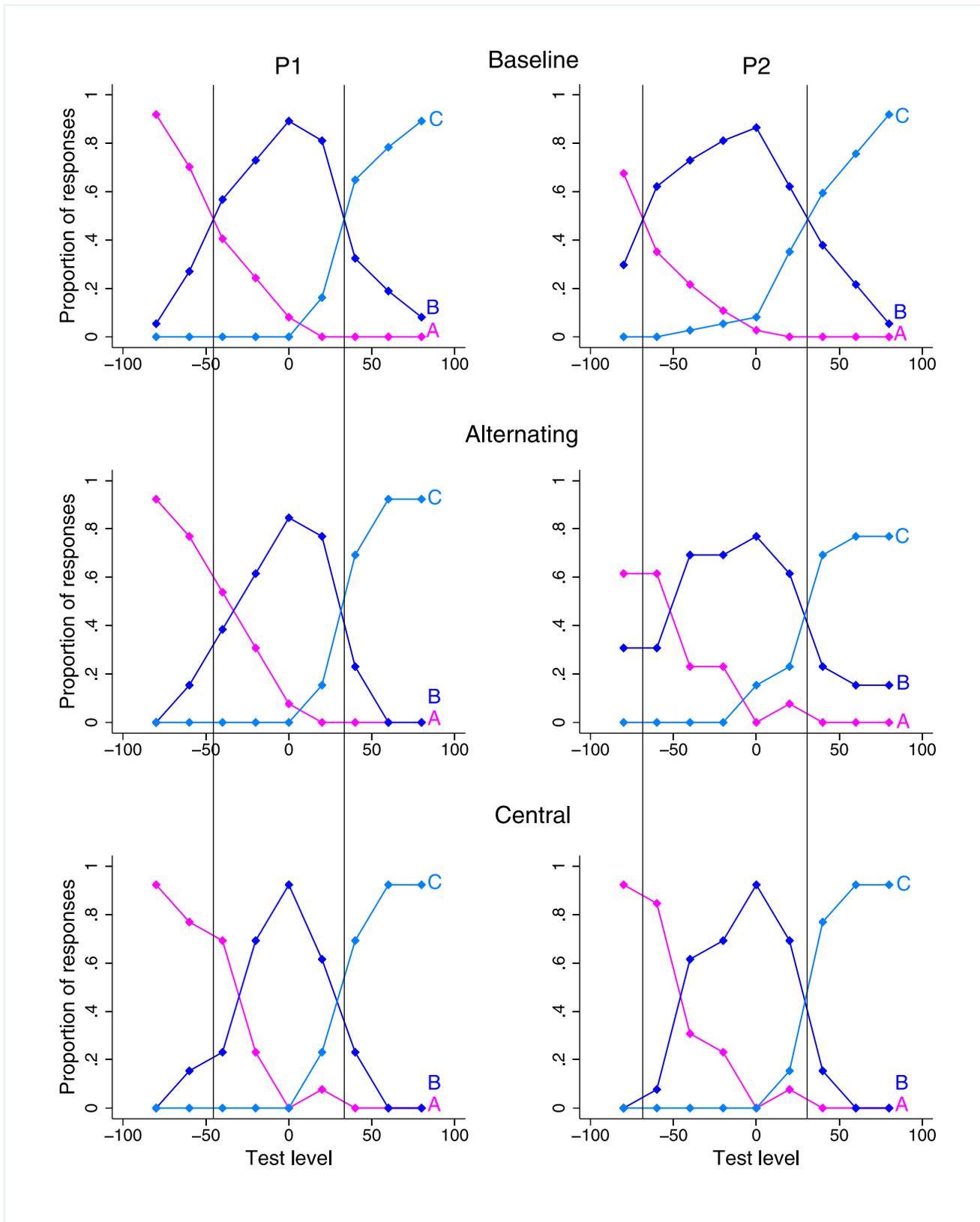


Figure 6. Examples of individual participants' data: Mean proportion of “A” (antifear, shown in magenta), “B” (average, shown in dark blue), and “C” (fear, shown in light blue) responses to each level of the test trajectory in the baseline, alternating adaptation, and central adaptation conditions for two participants (left and right columns). To aid comparison across adaptation conditions, we show vertical lines that indicate the test level of the A-B and B-C crossing points in the baseline condition.

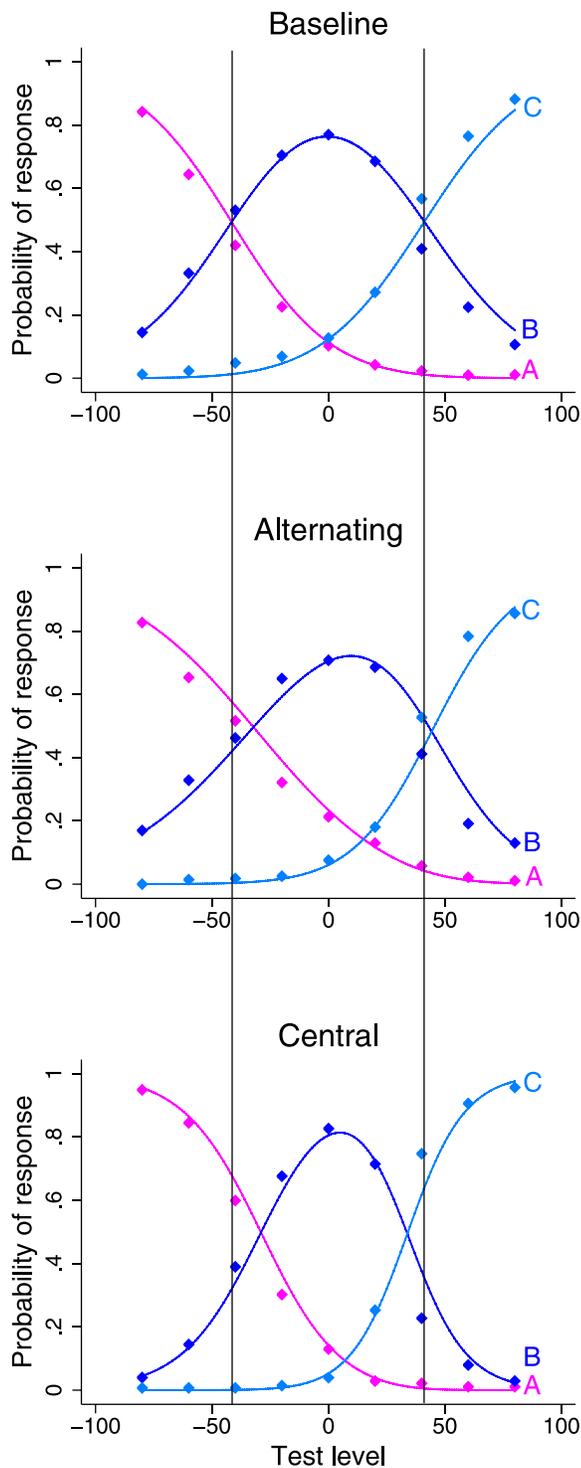


Figure 7. Response curves estimated by the mixed logit model. Curves show the probability of “A” (antifear, shown in magenta), “B” (average, shown in dark blue), and “C” (fear, shown in light blue) responses to each level of the test trajectory in the baseline, alternating adaptation, and central adaptation conditions. To aid comparison across adaptation conditions, we show vertical lines that indicate the test level of the A-B and B-C crossing points in the baseline condition. Points show mean proportion of responses from the group data as plotted in Figure 5.

the analysis makes it suitable for data such as ours, in which multiple observations are taken from each participant. The resulting functions are presented in Figure 7.

The points at which these functions crossed were determined for each adaptation condition. These crossing points represent the test level at which participants are equally likely to label an expression “B” or “A” and the test level at which participants are equally likely to label an expression “B” or “C.” We compared the distance between these crossing points across the different adaptation conditions. The distance between the crossing points indicates the perceived central range: the range of test levels that participants tended to label as “B” more than “A” or “C.” As can be seen in Figure 7, the central range of our fit functions narrowed after alternating adaptation compared to baseline and also narrowed after central adaptation compared to baseline. We estimated the standard error of these shifts, allowing us to test their significance, using the `nlcom` function of the STATA data analysis package (StataCorp, 2013), which uses the delta method (Oehlert, 1992). The interthreshold distance was significantly narrower than it was at baseline in both the alternating and central conditions (Table 1). Importantly, the range changed in the same direction in both conditions, supporting the opponent-coding model.

Our second aim was to investigate the possible functional benefit of expression adaptation. The narrowing of the central range observed above indicates that participants were more likely to identify fear and antifear around the average in both adaptation conditions. However, there was a greater narrowing of the central range for central adaptation compared to alternating adaptation,  $z = -4.96$ ,  $p < 0.001$ ,  $d = 0.74$ , indicating that central adaptation increased identification of subtle expressions more than alternating adaptation did. To further examine the effect of adaptation on participants’ identification of fear and antifear, we compared their responses to the strongest test expressions,  $-80\%$  antifear and  $80\%$  fear, between adaptation and baseline. As can be seen in Figure 7, central adaptation made the  $-80\%$  expression more likely to be labeled “A” (antifear),  $z = 9.60$ ,  $p < 0.001$ ,  $d = 1.64$ , and the  $80\%$  expression more likely to be labeled “C” (fear),  $z = 10.60$ ,  $p < 0.001$ ,  $d = 1.82$ . Alternating adaptation did not significantly change the likelihood of identifying the  $-80\%$  or  $80\%$  expressions as “A” and “C,” respectively,  $z = -1.07$ ,  $p = 0.283$ ,  $d = 0.21$  and  $z = 1.71$ ,  $p = 0.088$ ,  $d = 0.28$ , respectively. Therefore, although both adaptation conditions increased identification of subtle expressions around the average, only adaptation to the average expression increased identification of the stronger expressions at the ends of the trajectory. Again, this suggests that

Adaptation condition	Central range		Change in central range from baseline				
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>d</i> <sup>a</sup>
Baseline	82.74	2.26					
Alternating	76.55	2.88	−6.19	2.52	−2.45	0.014	0.35
Central	62.83	1.94	−19.09	2.10	−9.47	<0.001	1.28

Table 1. The range of test levels judged as central (“B”) at baseline and in each adaptation condition and tests of the significance of the changes in this range from the baseline condition. *Notes:* Range is shown in units of adaptor strength percentage. <sup>a</sup> Cohen’s *d* was calculated here by dividing the difference between the baseline and postadaptation values by the pooled standard deviations of the baseline and postadaptation values (Dunlap et al., 1996). Standard deviations were produced by multiplying the approximated standard errors by the square root of *N*.

adapting to the average had a greater effect on identification of expressions than adapting to the extremes of the trajectory.

Alternating adaptation also caused the functions describing participant responses to shift towards the “fear” end of the trajectory, indicating a bias to see faces as less fearful in this condition. Both the A-B and B-C thresholds shifted significantly toward “fear” compared to baseline,  $z = 4.90, p < 0.001, d = 1.07$  and  $z = 2.02, p = 0.043, d = 0.44$ , respectively. In the central condition, both the A-B and B-C thresholds shifted significantly inward, toward 0%,  $z = 8.82, p < 0.001, d = 1.76$  and  $z = 4.94, p < 0.001, d = 1.13$ , respectively, indicating that there was no overall bias away from “fear” responses in this condition. These results indicate that there may be an imbalance in the size of the aftereffects produced by fear and antifear.

In the analyses above, we compared the effects of adaptation between the central and alternating conditions. It is important to be sure that participants attended equally to the adaptors in both conditions as better-attended adaptors tend to produce larger aftereffects (Rhodes et al., 2011). To check that participants were not attending more to the adaptor in the central condition than in the alternating condition, we looked at performance in the change detection task that took place during the long period of initial adaptation. We calculated the proportion of eye changes and lip changes that were correctly identified for each of the three adaptors (0%, central condition, and −100% and 100%, alternating condition) (Table 2). A repeated-measures ANOVA revealed no significant main effect

Adaptor	Eye changes		Lip changes		Overall	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
0%	.76	.04	.63	.04	.70	.04
−100%	.72	.04	.71	.04	.71	.03
100%	.75	.04	.66	.04	.71	.04

Table 2. The proportion of eye and lip changes correctly identified for each adaptor (−100% antifear, 0% average, and 100% fear) and overall proportion of changes detected for each adaptor during the change detection task.

of adaptor,  $F(2, 44) = 0.11, p = 0.896, \eta^2_p = .01$ ; a significant main effect of feature,  $F(1, 22) = 4.28, p = 0.050, \eta^2_p = .16$ ; and no significant interaction,  $F(2, 44) = 1.59, p = 0.215, \eta^2_p = .07$ . Thus, although participants were significantly better at detecting eye changes than lip changes, there was no difference in their performance across the three adaptors. This finding indicates that the differences we found between the central and alternating adaptation conditions were not due to differences in attention.

## Discussion

We investigated whether a central-channel model, rather than a pure opponent model, is necessary to explain the coding of facial expression. To do so, we compared the effects of adapting to either alternating extreme adaptors (fear and antifear) or a central adaptor (average) on the range of expressions seen as central on a fear–antifear expression trajectory. We found a significant inward shift in the central range for both the alternating and central adaptation conditions relative to baseline. This pattern of results is not predicted by a central-channel model. It can, however, be accommodated by an opponent-coded model.

Our findings agree with several other studies that have found evidence for the opponent coding of facial expression (Burton et al., 2013; Skinner & Benton, 2009, 2010). These previous studies have used the near–far paradigm, which compares the sizes of the aftereffects produced by weak and strong antiexpressions. Stronger antiexpressions produce larger aftereffects, indicating norm-based coding. However, this paradigm does not necessarily distinguish between an opponent-coding model and certain multichannel models. The present study used a different paradigm not previously applied to facial expression and provides converging evidence in support of the opponent coding of expression.

As well as perceptual adaptation, changes in decision-making may have contributed to the shifts in responses that we see here. By the nature of this

paradigm, the adaptors are also the exemplars of the three expression categories. Seeing these expressions during the adaptation blocks may have provided a reference that helped to reduce uncertainty in participants' choices or changed the criteria by which they made those choices. In future, this potential referencing effect could be reduced (although not eliminated) by using different identities for adaptation and testing. However, there is evidence that a portion of the expression aftereffect is identity-dependent (Skinner & Benton, 2012), so future research taking this measure would need to take into account an expected decrease in the size of the aftereffects.

Our results may indicate a functional role of adaptation. The narrowing of the central range in both adaptation conditions relative to baseline indicates an increase in identification of fear and antifear after adapting to faces from the trajectory. Adaptation may help to calibrate perceptual resources to recently seen stimuli, improving discrimination around the range of expressions where it is currently needed. This functional benefit of adaptation has previously been supported in the perception of other facial attributes, such as identity (Rhodes, Watson, Jeffery, & Clifford, 2010), gender (Yang, Shen, Chen, & Fang, 2011), and viewpoint (Chen, Yang, Wang, & Fang, 2010), but has not previously been established for facial expression.

We should be cautious, however, about interpreting increases in expression identification following adaptation as an improvement in sensitivity. Again, these effects could also be explained by changes in participants' decision criteria in the adaptation conditions. Viewing the adaptors may give participants a reference point against which to judge the test faces, altering their responses. For instance, if a participant tended to respond "B" when uncertain, reducing that uncertainty by providing a reference would improve the identification rates for fear and antifear expressions. Any potential functional benefit of expression adaptation should be investigated in future research using a measure of sensitivity that is less affected by response biases, such as an odd-man-out task (O'Mahony, 1995).

It is interesting to consider that we were able to produce aftereffects in our participants after adapting them to the average expression, which is an approximation of the norm. In norm-based models, aftereffects are typically explained as the result of shifts in the position of the norm produced by adaptation. Adaptation to the norm itself should not produce such a shift and so might not be predicted to produce an aftereffect. However, in the opponent-coded instantiation of norm-based coding supported by our findings, changing the position of the norm may not be the only way to produce an aftereffect. Adaptation may also change the slope of the tuning functions of the opponent pools so

that the relative response rate of the two pools to a given stimulus may be altered even when the norm, the point at which they respond equally, remains the same. This kind of change in response can be visualized as a distortion of expression space that stretches the space around the norm, making expressions around that point appear more distinctive.

It should be noted that, following Calder et al. (2008) and Lawson et al. (2009, 2011), our initial predictions are based on the assumption that adaptation suppresses the neural channels stimulated by the adaptor. If adaptation also affects the shape of the response curves, it would make predictions about the result of adaptation in each model more complex. In the case of the opponent-coding model, the simplest situation is one in which adaptation in both the central and alternating conditions affects the shape of the response curves in the same way. If this is the case, our initial prediction (that both alternating and central adaptation will have the same effect on the size of the central range) remains the same. It is also possible (but less plausible) that adaptation in one condition might steepen the response curves, and adaptation in the other condition might flatten them. If this were the case, the opponent-coding model might be able to accommodate opposing changes in the size of the central range. The central-channel model is even more complex as there are three or more neural channels to consider, and the response curve of the central channel may be affected by adaptation differently to that of the outer channels. Computer simulation of the models may be helpful for determining the predicted effects of this adaptation paradigm as we vary our assumptions (cf. Ross, Deroche, & Palmeri, 2013, for an example of how this might be approached).

We chose an average expression as our approximation of the norm as has been done in several other expression aftereffect studies (Burton et al., 2013; Cook et al., 2011; Skinner & Benton, 2010, 2012). We chose this expression because the norm represents the central tendency of previously seen faces (Valentine, 1991). However, other studies have instead used a neutral expression as their norm (Juricevic & Webster, 2012; Rutherford, Chattha, & Krysko, 2008; Rutherford, Troubridge, & Walsh, 2011). The neutral expression represents the face at rest, which could be argued would make it a good "unexpressive" candidate for the norm, analogous to white in color space (Juricevic & Webster, 2012). However, there is evidence that the neutral face is not actually perceived as expressively neutral. In an implicit emotion evaluation task, Lee, Kang, Park, Kim, and An (2008) found that participants responded to neutral stimuli in the same way they did negative stimuli. Additionally, in studies that map the perceptual organization of emotion based on participants' judgments of expressions, neutral expressions are often

located away from the center of the space, nearer to other prototypical emotion expressions (Bimler & Kirkland, 2001; Gao, Maurer, & Nishimura, 2010; Russell & Bullock, 1985, 1986). An average expression may therefore be the more appropriate choice.

We found an overall bias toward “antifear” responses in the alternating adaptation condition. This bias suggests an imbalance in the amount of adaptation produced by the 100% and –100% expressions perhaps because the fear expression is more familiar or more attention-grabbing than antifear. Better-attended adaptors tend to produce larger aftereffects (Rhodes et al., 2011). However, there was no difference in change detection performance between the two adaptors, so any difference in attention was not large enough to affect performance this task. The shape and textural information of the –100% antifear expression differ from the average to the same extent as those of the 100% fear expression. Nevertheless, it is possible that 100% fear is more perceptually dissimilar to the average expression than –100% antifear, which would result in stronger adaptation to the fear end of the trajectory (Robbins et al., 2007). This imbalance is not a problem for the interpretation of our data, however, because we still found a significant reduction of the range of expressions seen as central following alternating adaptation as well as following central adaptation.

Although we are using a fear–antifear trajectory here to draw conclusions about the coding systems that underlie expression perception, we do not assume that fear–antifear is an explicitly coded dimension of expression space or that there are specific antifear detectors that are affected by adaptation to antifear. Rather, as in previous studies (Burton et al., 2013; Skinner & Benton, 2010, 2012), we assume that adaptation along the fear–antifear trajectory adapts a number of underlying, expression-relevant dimensions. These dimensions might relate to the positions of particular features or muscle groups or could describe key components of the statistical variation found in facial postures (Cook et al., 2011). By studying participants’ judgments of the composite fear–antifear trajectory, we can indirectly observe the adaptation of those underlying dimensions.

Our method of data analysis differed from that of Lawson et al. (2009, 2011) and Calder et al. (2008). In those studies, five test levels were used (e.g., 10° left, 5° left, direct, 5° right and 10° right for gaze direction), and changes to the central range were approximated by calculating changes in the proportion of expressions judged to be central at levels immediately to either side of the center of the trajectory (e.g., 5° left and 5° right in the above example). An increase in the proportion judged to be central at these levels indicated a widening of the central range; a decrease at these levels indicated a narrowing of the central range. As the alternating

condition in our study caused an overall bias in responses toward “fear,” this simple analysis was not appropriate for our data. Fitting the mixed multinomial logit model allowed us to more directly test the predicted changes in the central range and to quantify the size of these changes. It should be noted that for mixed logit models, there is no intuitive measure of goodness of fit (Train, 2003). The log likelihood ratio indicates that our model fit better than the unconstrained model, but, unlike an  $R^2$ , this measure cannot be interpreted in terms of the proportion of variance explained. However, the closeness of the fitted functions (Figure 7) to the pattern of mean proportions of responses (Figure 5) confirms that the model is an appropriate representation of the pattern of the data.

In the present study, we used a single expression trajectory, running from fear to antifear. Testing one trajectory allowed us to maximize the power of our analysis. Expressions are modeled as being coded in a single, multidimensional expression space, so the coding system that applies to the fear–antifear trajectory should be common to other expression trajectories. However, in future, it may be useful to test other expression trajectories to ensure that our findings generalize across multiple expressions. In particular, use of a trajectory that runs between two unfamiliar expressions (such as the trajectories defined by primary components analysis in Cook et al., 2011) may prevent the asymmetric adaptation created by using a familiar expression and unfamiliar antiexpression as the endpoints of the test trajectory.

Future research should seek converging evidence that expression is opponent coded by applying other paradigms that can distinguish between opponent and multichannel models to expression stimuli. One such paradigm, used previously to examine the coding of gender in faces (Pond et al., 2013), involves adapting participants to a very wide range of adaptor strengths. This approach improves on the near–far paradigm by using adaptors that extend beyond the natural range of facial variation, where we are more likely to capture differences between opponent and multichannel coding. If, as suggested here, facial expression is opponent coded, we would expect expression aftereffects to increase with increasing adaptor strength across the natural range of variation and then either continue to increase or asymptote beyond that range.

Our present findings support an opponent-coding model of facial expression rather than a central-channel model. They are in line with previous studies that have found evidence supporting the opponent coding of facial expression using the near–far paradigm (Burton et al., 2013; Skinner & Benton, 2010, 2012). We also found that adaptation to a range of expressions, and particularly the central tendency of that range, resulted in increased identification of the fear and antifear

expressions, possibly indicating a functional benefit of adaptation for expression perception.

*Keywords:* facial expression, face adaptation, face aftereffects, expression aftereffects, opponent coding

## Acknowledgments

This research was supported the UK Medical Research Council under project code MC-A060-5PQ50 (Andrew J. Calder), the Australian Research Council Centre of Excellence in Cognition and its Disorders (CE110001021), an ARC Professorial Fellowship to Rhodes (DP0877379), and an ARC Discovery Outstanding Researcher Award to Rhodes (DP130102300). We thank Andrew Skinner and Chris Benton for providing the stimuli, Michael Burton for assistance with data analysis, and Colin Clifford for helpful discussions. Ethical approval was granted by the Human Research Ethics Committee of the University of Western Australia. This paper is dedicated to the memory of Andy Calder (1965–2013).

Commercial relationships: none.

Corresponding author: Nichola Burton.

Email: nichola.burton@uwa.edu.au.

Address: ARC Centre of Excellence in Cognition and its Disorders, School of Psychology, The University of Western Australia, Crawley, WA, Australia.

## Footnote

<sup>1</sup> In the first session, a computer error caused one participant to repeat the no-feedback training phase 46 times before finishing training. This participant has been left out of the training descriptives for the first session. In the second session, the participant only required five repetitions to finish the training.

## References

Armann, R., Jeffery, L., Calder, A. J., Bühlhoff, I., & Rhodes, G. (2011). Race-specific norms for coding face identity and a functional role for norms. *Journal of Vision*, *11*(13):9, 1–14, <http://www.journalofvision.org/content/11/13/9>, doi:10.1167/11.13.9. [PubMed] [Article]

Bimler, D., & Kirkland, J. (2001). Categorical perception of facial expressions of emotion: Evidence from multidimensional scaling. *Cognition & Emotion*, *15*(5), 633–658.

Blakemore, C., & Sutton, P. (1969). Size adaptation: A new aftereffect. *Science*, *162*, 245–247.

Burton, N., Jeffery, L., Skinner, A. L., Benton, C. P., & Rhodes, G. (2013). Nine-year-old children use norm-based coding to visually represent facial expression. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1261–1269.

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*(2), 244–261.

Chen, J., Yang, H., Wang, A., & Fang, F. (2010). Perceptual consequences of face viewpoint adaptation: Face viewpoint aftereffect, changes of differential sensitivity to face view, and their relationship. *Journal of Vision*, *10*(3):12, 1–11, <http://www.journalofvision.org/content/10/3/12>, doi:10.1167/10.3.12. [PubMed] [Article]

Clifford, C. W. G. (2002). Perceptual adaptation: Motion parallels orientation. *Trends in Cognitive Sciences*, *6*(3), 136–143.

Clifford, C. W. G., & Rhodes, G. (Eds.). (2005). *Fitting the mind to the world: Adaptation and after-effects in high-level vision*. Oxford: Oxford University Press.

Clifford, C. W. G., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*, 1705–1710.

Cook, R., Matei, M., & Johnston, A. (2011). Exploring expression space: Adaptation to orthogonal and anti-expressions. *Journal of Vision*, *11*(4):2, 1–9, <http://www.journalofvision.org/content/11/4/2>, doi:10.1167/11.4.2. [PubMed] [Article]

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated-measures designs. *Psychological Methods*, *1*(2), 170–177.

Gao, X., Maurer, D., & Nishimura, M. (2010). Similarities and differences in the perceptual structure of facial expressions of children and adults. *Journal of Experimental Child Psychology*, *105*(1–2), 98–115.

Juricevic, I., & Webster, M. A. (2012). Selectivity of face aftereffects for expressions and anti-expressions. *Frontiers in Psychology*, *3*(4), 1–10.

Lawson, R. P., Clifford, C. W. G., & Calder, A. J. (2009). About turn: The visual representation of human body orientation revealed by adaptation. *Psychological Science*, *20*, 363–371.

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(9):17, 1–17, <http://www.journalofvision.org/content/11/9/17>, doi:10.1167/11.9.17. [PubMed] [Article]
- Lee, E., Kang, J. I., Park, I. H., Kim, J.-J., & An, S. K. (2008). Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Research*, 157(1–3), 77–85.
- Maddess, T., McCourt, M. E., Blakeslee, B., & Cunningham, R. B. (1988). Factors governing the adaptation of cells in area-17 of the cat visual cortex. *Biological Cybernetics*, 59, 229–236.
- Movshon, J. A., & Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature*, 278, 850–852.
- O'Mahony, M. (1995). Who told you the triangle test was simple? *Food Quality and Preference*, 6(4), 227–238.
- Oehlert, G. W. (1992). A note on the delta method. *The American Statistician*, 46(1), 27–29.
- Pond, S., Kloth, N., McKone, E., Jeffery, L., Irons, J., & Rhodes, G. (2013). Aftereffects support opponent coding of face gender. *Journal of Vision*, 13(14):16, 1–19, <http://www.journalofvision.org/content/13/14/16>, doi:10.1167/13.14.16. [PubMed] [Article]
- Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, 46, 2977–2987.
- Rhodes, G., Jeffery, L., Evangelista, E., Ewing, L., Peters, M., & Taylor, L. (2011). Enhanced attention amplifies face adaptation. *Vision Research*, 51, 1811–1819.
- Rhodes, G., & Leopold, D. A. (2011). Adaptive norm-based coding of face identity. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 263–286). Oxford: Oxford University Press.
- Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., & Clifford, C. W. G. (2005). Adaptation and face perception: How aftereffects implicate norm-based coding of faces. In C. W. G. Clifford & G. Rhodes (Eds.), *Fitting the mind to the world: Adaptation and after-effects in high-level vision* (pp. 213–240). Oxford: Oxford University Press.
- Rhodes, G., Watson, T. L., Jeffery, L., & Clifford, C. W. G. (2010). Perceptual adaptation helps us identify faces. *Vision Research*, 50, 963–968.
- Robbins, R., McKone, E., & Edwards, M. (2007). Aftereffects for face attributes with different natural variability: Adapter position effects and neural models. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 570–592.
- Ross, D. A., Deroche, M., & Palmeri, T. J. (2013). Not just the norm: Exemplar-based models also predict face aftereffects. *Psychonomic Bulletin & Review*, 21(1), 47–70.
- Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, 48(5), 1290.
- Russell, J. A., & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology*, 22(1), 97.
- Rutherford, M., Chattha, H. M., & Krysko, K. M. (2008). The use of aftereffects in the study of relationships among emotion categories. *Journal of Experimental Psychology: Human Perception and Performance*, 34(1), 27–40.
- Rutherford, M., Troubridge, E., & Walsh, J. (2011). Visual afterimages of emotional faces in high functioning autism. *Journal of Autism and Developmental Disorders*, 42(2), 221–229.
- Skinner, A. L., & Benton, C. P. (2009). Adapting to anti-expressions: A journey through expression space. *Journal of Vision*, 9(8):523, <http://www.journalofvision.org/content/9/8/523>, doi:10.1167/9.8.523 [Abstract]
- Skinner, A. L., & Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychological Science*, 21(9), 1248–1253.
- Skinner, A. L., & Benton, C. P. (2012). The expressions of strangers: Our identity-independent representation of facial expression. *Journal of Vision*, 12(2):12, 1–13, <http://www.journalofvision.org/content/12/2/12>, doi:10.1167/12.2.12. [PubMed] [Article]
- StataCorp. (2013). *Stata statistical software: Release 13*. College Station, TX: StataCorp LP.
- Thompson, P., & Burr, D. (2009). Visual aftereffects. *Current Biology*, 19(1), R11–R14.
- Train, K. E. (2003). *Discrete choice with simulation*. Cambridge, UK: Cambridge University Press.
- Tsao, D. Y., & Freiwald, W. A. (2006). What's so special about the average face? *Trends in Cognitive Sciences*, 10, 391–393.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology*, 43A, 161–240.
- Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society: Biological Sciences*, 366, 1702–1725.
- Yang, H., Shen, J., Chen, J., & Fang, F. (2011). Face adaptation improves gender discrimination. *Vision Research*, 51(1), 105–110.