Adaptation to antifaces and the perception of correct famous identity in an average face

Running head: Adaptation to antifaces

Anthony C Little[1*], Peter JB Hancock[1], Lisa M DeBruine[2], Benedict C Jones[2]
[1] School of Natural Sciences, University of Stirling, Stirling, FK9 4LA, Scotland, UK
[2] School of Psychology, University of Aberdeen, Aberdeen, AB24 2UB, Scotland, UK

***Correspondence:**
Dr Anthony C. Little,
School of Natural Sciences,
University of Stirling,
Stirling, FK9 4LA, Scotland, UK
Email: anthony.little@stir.ac.uk
Phone: (+44) 01786 467651
Fax: (+44) 01786 467641

Abstract

Previous experiments have examined exposure to anti-identities (faces that possess traits opposite to an identity through a population average), finding that exposure to antifaces enhances recognition of the plus-identity images. Here we examine adaptation to antifaces using famous female celebrities. We demonstrate: that exposure to a color and shape transformed antiface of a celebrity increases the likelihood of perceiving the identity from which the antiface was manufactured in a composite face and that the effect shows size invariance (Experiment 1), equivalent effects are seen in internet and laboratory based studies (Experiment 2), adaptation to shape-only antifaces has stronger effects on identity recognition than adaptation to color-only antifaces (Experiment 3), and exposure to male versions of the antifaces does not influence the perception of female's faces (Experiment 4). Across these studies we found an effect of order where aftereffects were more pronounced in early than later trials. Overall, our studies delineate several aspects of identity aftereffects and support the proposal that identity is coded relative to other faces with special reference to a relatively sex-specific mean face representation.

Key words:

Aftereffects; adaptation; recognition; experience; face processing; prototypes; categories

Introduction

For each class of stimuli that the human visual system encounters, it may develop an individual representation, or prototype, made up of an average of the characteristics of all the different stimuli of that type that have been seen (Valentine, 1991;Enquist and Arak, 1994;Johnstone, 1994;Giese and Leopold, 2005;Loffler et al., 2005). Computer modeling has revealed that algorithms trained to discriminate different stimuli produce the strongest responses to stimuli that represent the average of the training set, even though this average was not previously encountered (Enquist and Arak, 1994;Johnstone, 1994). These findings have been interpreted as evidence that prototype formation is a property of learning to recognize different stimuli as members of a class (Enquist and Arak, 1994;Johnstone, 1994). Learning studies have examined how categorical perception develops using abstract stimuli. In classic studies it has been shown that exposure to different dot patterns with particular configurations results in abstraction so that the average of each of the patterns, while never previously seen, is recognized as belonging to the set of patterns from which it was derived (Posner and Keele, 1968). These results were originally taken as evidence for prototype formation, but there is much debate about whether it does represent abstraction of a prototype or whether people store individual exemplars and use these to determine category (Nosofsky and Zaki, 2002;Smith and Minda, 2002;Ashby and Maddox, 2005).

Faces have been the focus of much research regarding recognition and possible prototype formation. The notion of a multidimensional ‘face space’ has proved extremely useful in helping to understand how faces are mentally represented (Valentine, 1991). In face space, each individual face is a point in a space with a theoretical average of all faces at the center. Individual faces lie on trajectories or vectors and caricatures move them outwards, away from the centre of the distribution, while anti-caricatures move the face towards the centre. Caricatures move away from average and so make faces more distinctive and recognized more quickly than veridical original identities (Rhodes et al., 1987;Benson and Perrett, 1994). Such studies were interpreted as evidence that the average face played a special role representing face identity (e.g., Benson and Perrett, 1994), a ‘prototype’ model of face coding. Many researchers have thought that a prototype face could function as a norm that anchors coding of identity (Valentine, 1991;Leopold et al., 2001). Of course, this idea has been debated, with some researchers rejecting the view that average face plays a special role and supporting instead the view that faces may be coded more directly as veridical representations of individuals or exemplars, an ‘exemplar’ model (Valentine, 1991).

Studies of face adaptation to identity have been taken as evidence for a prototype model. Exposure to faces (i.e. adaptation) biases subsequent perceptions of novel faces by causing faces similar to those initially viewed to appear more prototypical (i.e. normal) than they would otherwise be perceived, reflecting the recalibration of a prototype in light of recent visual experience (Leopold et al., 2001;Rhodes et al., 2001;Rhodes et al., 2003;Rhodes et al., 2004;Webster et al., 2004;Leopold et al., 2005;see Webster and MacLeod, 2011 for review). We note that in these studies it is also plausible that a population of exemplars becomes updated. In adaptation studies, for example, adaptation (exposure) to faces with contracted features causes novel faces with contracted features to

be perceived as more normal than prior to this exposure (Rhodes et al., 2003;Rhodes et al., 2004). Analogous visual aftereffects have been observed following exposure to faces varying in identity (Leopold et al., 2001;Rhodes et al., 2001), ethnicity (Webster et al., 2004), sex (Rhodes et al., 2004;Webster et al., 2004), expression (Webster et al., 2004) and sexual dimorphism (Little et al., 2005;Buckingham et al., 2006).

While these findings might be consistent with an exemplar view, an adaptation study by Leopold et al. (2001) has shown evidence more in favor of prototype views. The researchers made use of 'antifaces' for particular identities. Antifaces lie on the same vector as the original identity but exist in the other side of the average face and therefore have the appearance of a face possessing the opposite traits of the original face. For example, an antiface for an identity with a bigger-than-average nose would be a face that had a smaller-than-average nose. To compare, a caricature would have a larger nose than the original and an anti-caricature would have a nose somewhere between the original and average. Leopold et al. (2001) found that identification of a particular identity (e.g., Jim) was made easier by adapting to the antiface (AntiJim), but not to other antifaces. This finding has been replicated in later studies using similar methods (Leopold et al., 2005) and using entirely computer generated images faces (Anderson and Wilson, 2005). As adaptation biases perception along an identity trajectory away from the adapting stimulus, the antiface effects in these studies strongly suggest the average face defines such a trajectory. Importantly for this interpretation, the aftereffect must be selective for computationally opposite identities and not be a generalized contrast effect. This is possible if the non-opposite adapting faces were more similar to the test or probe faces than were the opposite adapting faces and so adaptation to non-opposite faces could lead to a general facilitation in any direction of face space rather than along a particular direction due to differences in perceptual contrast (Rhodes and Jeffery, 2006). Rhodes and Jeffery (2006) measured the face identity aftereffect for computationally opposite and non-opposite adapt-test pairs that had been matched on perceived dissimilarity to control for this confound and found similar results to previous studies, supporting a special role for the average face. Interestingly, the authors also suggest a model of coding that allows individuating information to be coded as deviations from average or prototype without an explicitly mentally represented average (Rhodes and Jeffery, 2006).

Recent neuroimaging and single-cell recording studies have also supported a prototype-referenced model of face coding (Giese and Leopold, 2005;Loffler et al., 2005). For example, neuroimaging of humans during adaptation to facial identities has shown that specific neural populations respond to faces falling along specific identity axes away from an average (Loffler et al., 2005) and similar results are seen whereby neurons in the macaque brain represent deviations from average (Giese and Leopold, 2005). Further experimental evidence comes from a recent study by Ross, Hancock and Lewis (Ross et al., In Press), which shows a decreased sensitivity to changes along the axis between a face and the average (caricature/anti-caricature) compared with equivalent sized changes in a different direction (lateral caricatures). This suggests that there is something special about the position of the average face and the axis between it and a given face.

Previous studies have examined the time-course of identity adaptation to elucidate the phenomenon, such as the effect of adaptation duration (Leopold et al., 2005). Here we examine effects of the nature of the stimuli used. The current study investigated whether adaptation to antifaces made from famous female celebrities would result in accurate perception of identity in a neutral face (Experiment 1A) and whether this effect held true across size transformations (Experiment 1B). We additionally compared intent and laboratory based experiments (Experiment 2) and also tested for independent contributions of color and shape information to identity aftereffects (Experiment 3A and 3B). It has been shown that adaptation to eye-spacing, face shape, and sexual dimorphism can influence the perception of male and female faces in different directions simultaneously for judgments of both normality and attractiveness (Little et al., 2005), suggesting that faces are coded relative to a sex-specific norm (see also Bestelmeyer et al., 2008;Jaquet and Rhodes, 2008). Face identity aftereffects for newly learned faces , as measured by recognition after exposure to antifaces, were significantly larger for pairs using an opposite which was of the same sex than an androgynous average suggesting identity is coded relative to a sex-specific norm (Rhodes et al., 2011). Thus, we also examined whether adaptation to male versions of female antifaces can influence the perception of female faces (Experiment 4). Across all of the experiments we examined the effects of trial order because pilot studies conducted by the first and second author suggested adaptation to antifaces had greatest influence on recognition in early trials.

Previous studies of identity adaptation have tested the threshold for recognition of a newly-learned face by presenting a series of anti-caricatures of varying strengths. Using this technique, for example, Rhodes and Jeffery (2006) found a change in threshold (50% correct identification) from 55.4% of the distance from average to target face to 43.5%. Two key differences here are that we are using famous faces, for which most of our participants should have a well-established representation, and that we present only the average face for testing. This type of test therefore has no inherent cues to the identity of the target and is a stronger test of the ability of adaptation to produce identification. Previous studies addressing identity adaptation use newly learned faces, faces learned under the same experimental conditions as the adaptation tests are carried out. By using famous faces we bypass any potential effects of learning/test congruence and examine a more natural form of familiarity. Such demonstrations are clearly important if identity adaptation effects are to have real world validity.

Experiment 1A and 1B

We examined the influence of adaptation to antifaces for female celebrity images and subsequent perception of identity in a neutral probe (Experiment 1A/B). Previous studies address newly learned faces and here we examine faces with which participants are likely to have longer term familiarity. Following previous studies, we examined the influence of changing size in the probe stimuli (Experiment 1B) to examine the extent such aftereffects reflect adaptation of higher-level neural mechanisms (Leopold et al., 2001). Adaptation was tested by presenting participants with an ‘antiface’, the opposite shape and color of a particular identity for 6 seconds, followed by a neutral average face for 1 second. Participants were then presented with the names of two identities, one correct and

one incorrect, and asked to indicate which name the second face they saw most resembled. Following previous studies, we predicted that exposure to antifaces would result in an aftereffect such that a neutral face would resemble the real identity, leading to correct identification.

Participants

Participants in Experiment 1A were 15 women and 17 men (mean age = 23.0, SD=4.5). Participants in Experiment 1B were 30 women and 9 men (mean age = 33.8, SD = 13.2). Different participants took part in each experiment. All participants were volunteers and were selected for being between the ages of 16-65 and being UK or US nationals to help ensure familiarity with the celebrity faces. The study was run over the internet and participants were recruited via a link from a research based website (www.alittlelab.com).

Stimuli

All stimuli were constructed using established (Perrett et al., 1998;Little et al., 2001;Little et al., 2005) techniques for averaging and manipulating the appearance of face images in an objective, systematic manner (for technical details including mathematical algorithms see Benson and Perrett, 1993;Rowland and Perrett, 1995;Tiddeman et al., 2001).

First we created 20 composite female celebrity images. Ten images each of 20 female celebrities were collected and a composite image created for each. Composites were made by manually marking 179 feature landmark points on all faces delineating the main features (e.g., points outline, eyes, nose, and mouth) and the outline of each face (e.g., jaw line, hair line). The average location of each point on the faces for each individual composite was then calculated. The features of the individual faces were then morphed to the relevant average shape before superimposing the images to produce a photographic quality result average in both shape and color. Examples of individual composites can be seen in Figure 1. We created an average female celebrity image to act as the centre of our feature space and this was done by averaging all 20 individual composite faces into a single image as described above. All images were symmetrized and standardized for size on interpupillary distance prior to transformation.

Antiface images were created by transforming the average female composite relative to a pair of face images specific to each identity: the individual celebrity composite image and the average female composite (see Figure 1). For example, using the difference between the composite Jennifer Aniston and the average composite we can compute a face that lies on the mirror opposite – an anti-100% face (See Figure 2: right hand panel). This procedure for manipulating faces through the mean is methodologically similar to that used in previous studies (Leopold et al., 2001).

Antifaces were used as adapting stimuli in the trials and the same average celebrity face was used as the probe in the trials. All images were resized to 360x496 pixels and an additional probe face was created at 70% of this size for use in Experiment 1B.

Figure 1 about here

Procedure

Participants first filled in a short questionnaire addressing their age, sex, and nationality. They were then presented with the following instruction: "In this study you will see faces and be asked to guess which celebrity is hidden within the images. You will be presented with a fixation image (a white cross on a black background) followed by a face. Please stare at this face. After a few seconds it will disappear and a new face will appear briefly. You will then be asked to name which celebrity the 2nd face looks like. The 2nd face will look very similar each time and you may feel like you do not know who it looks like. Please guess anyway. Guessing will move you onto the next trial."

Each trial consisted of a fixation image presented for 1 second, an adapting antiface image for 6 seconds, the probe trial composite female image for 1 second and finally the response image which asked participants to guess the celebrity (See Figure 2). Underneath the response image two celebrity names were presented as buttons. One name corresponded to the correct celebrity antiface and the other a random other name from the set. Names were paired such that, for example, names A and B when used as a pair were preceded once by anti-face A and once by anti-face B. Button order was randomized by side. Participants selected the name from the two alternatives which started the next trial. Trial order was randomized for each participant.

After the adaptation trials, participants were presented with the 20 original celebrity averages, for a recognition test. The recognition test composed of unmanipulated composite of each celebrity and their name for which participants were asked "do you recognize this face?" with boxes presented alongside marked "yes" and "no". This test was presented on a single page and the faces presented in alphabetical order of surname.

Figure 2 about here

Results

We calculated percent of correct answers counting correct answers for only those trials for which the participant subsequently stated they recognized both of the celebrities for the appropriate trial. We also used scores that included all the answers given irrespective of stated recognition. We used these latter measures to calculate the correlation for each participant between order of trial and correct answers and average accuracy for the first 10 trials versus the last 10 trials as measures of relative accuracy between the start and end of the test.

Recognition rates across 1A and 1B in the post-test were very high with an average of 94% (SD = 14.9) of faces used in the test being recognized in their unaltered form. All tests are presented 2-tailed.

Identity was recognized at rates greater than chance in both experiments. One-sample t-tests against chance (50%) revealed that participants showed correct recognition of identity in Experiment 1A ($t(31) = 2.75$, $p = .010$) and Experiment 1B ($t(38) = 2.37$, $p = .023$). An independent samples t-test revealed no significant difference in accuracy between Experiment 1A and 1B ($t(67) = 0.82$, $p = .415$). Means can be seen in Figure 3.

Figure 3 about here

Additional tests demonstrated that there was a weak effect of order such that accuracy was higher for earlier than later trials. One-sample t-tests against chance (0 = no correlation) revealed significant negative correlations between trial order and correct responses for both full size (mean r = -.12, SD = .28, $t(31) = 2.46$, $p = .019$) and 70% (mean r = -.08, SD = .23, $t(38) = 2.31$, $p = .027$) images. Paired samples t-test revealed that participants were more accurate for the first 10 trials than the second 10 trials for both the full-size images (1st 10: mean = 61.3%, SD = 19.8; 2nd 10: mean = 49.1%, SD = 17.1; $t(31) = 2.63$, $p = .013$) and the 70% size images (1st 10: mean = 57.7%, SD = 16.3; 2nd 10: mean = 51.0%, SD = 16.6; $t(38) = 1.86$, $p = .071$), though for the latter comparison the p-value was only approaching significance.

Discussion

Experiment 1A replicates identity aftereffects using famous faces rather than identities learned for the study. Experiment 1B also replicates scaling invariance of identity aftereffects using famous faces. As noted by others, it is possible that perception of correct identity after exposure to antifaces could be explained by several low-level aftereffects for orientation, spatial frequency and color (Leopold et al., 2001). Following Leopold et al. (2001), we also note that such low-level adaptation appears unlikely, as individuals were free to scan the adapting images, which continually varies the retinal locations of the various facial features and such effects were not significantly influenced by a decrease in the size of the probe image. Robustness to changes in size is most consistent with identity aftereffects being related to higher-level adaptation, as noted by several previous authors (e.g., Leopold et al., 2001).

We also found consistent effects of order of trial such that aftereffects resulted in accurate perceptions of identity for early trials and less so for later trials. This finding is considered further in the general discussion.

Experiment 2

Experiments 1A and 1B were conducted over the internet. To address the validity of web-based tests of adaptation we replicated Experiment 1A under laboratory conditions.

Participants

Participants in Experiment 2 were 13 women and 11 men (mean age = 27.5, SD=10.9). All participants were volunteers and were selected for being between the ages of 16-65

and being UK or US nationals to help ensure familiarity with the celebrity faces. The study was run under laboratory conditions.

Stimuli

Stimuli were identical to those used in Experiment 1A.

Procedure

The procedure was identical to Experiment 1A except that the experiment was taken under laboratory conditions. The laboratory testing took place on one of two identical computers in the same laboratory. Images were presented on a 32-bit colour 21 inch (1280x1024 pixels) LCD monitor. Stimuli subtended approximately 9 by 12.4◦ of visual angle when viewed by participants approximately 80 cm from the computer screen.

Results

Recognition rates for Experiment 2 in the post-test were high with an average of 98% (SD = 5.5) of faces used in the test being recognized in their unaltered form.

Identity was recognized at rates greater than chance in Experiment 2 and no difference in effect was found between Experiment 1A and 2. One-sample t-tests against chance (50%) revealed that participants showed correct recognition of identity in Experiment 2 ($t(23) = 2.50$, $p = .020$). An independent samples t-test revealed no significant difference in accuracy between Experiment 2 and Experiment 1A ($t(54) = 0.19$, $p = .954$). Means can be seen in Figure 4.

Figure 4 about here

An additional independent samples t-test revealed a significant difference in recognition rates between Experiment 2 and Experiment 1A ($t(54) = 2.08$, $p = .042$). Individuals recognized more of the celebrities under laboratory conditions, though this cannot influence accuracy as only trials in which participants recognized both celebrities were used to calculate accuracy.

Additional tests demonstrated that there was a weak effect of order such that accuracy was higher for earlier than later trials. A one-sample t-tests against chance (0 = no correlation) revealed a close to significant negative correlation between trial order and correct responses (mean r = -.10, SD = .25, $t(23) = 1.94$, $p = .065$). A paired samples t-test revealed that participants were more accurate for the first 10 trials than the second 10 trials (1st 10: mean = 63.3%, SD = 19.5; 2nd 10: mean = 52.1%, SD = 18.9; $t(23) = 2.30$, $p = .031$).

Discussion

Experiment 2 replicated the findings of Experiment 1A under laboratory conditions. While there was evidence that participants recognized more celebrities in the laboratory (90% vs. 98%) the size of the effect of adaptation on accuracy (correct recognition) was

almost identical. Likewise, order effects were of almost equivalent magnitude in Experiment 2 as in Experiment 1A. Experiment 2 then demonstrates equivalency in results based on web-based and laboratory based adaptation studies of this kind. We note that while individuals may intuitively feel that data collected in the absence of an experimenter may produce more variable results, there is growing evidence that adaptation effects seen under laboratory conditions are also seen in web based studies (Jones et al., 2008;Jones et al., 2010). Our experiment here demonstrates that the same effect is seen in Experiment 1 and 2 despite difference in recruitment, presence of experimenter, and variation in equipment.

Experiment 3A and 3B

Previous studies of identity aftereffects have simultaneously manipulated both shape and color information in face images when manufacturing adapting stimuli. By contrast, here we examined the independent contributions of shape and color in identity aftereffects.

Participants

Participants in Experiment 3A (shape only) were 15 women and 7 men (mean age = 27.5, SD=13.1). Participants in Experiment 3B (color only) were 24 women and 8 men (mean age = 29.4, SD = 9.2). Different participants took part in each experiment. Participants were selected as for Experiment 1A and 1B. The study was run over the internet and participants were recruited via a link from a research based website (www.alittlelab.com).

Stimuli

Stimuli were made in the same way as in Experiment 1A and 1B, but here transformations were anti-100% in shape only or anti-100% in color only. Example images can be seen in Figure 5.

Figure 5 about here

Procedure

The procedure was identical to that used in Experiment 1A. Participants were randomly allocated to a condition where they saw shape-only or color-only antifaces.

Results

We calculated percent of correct answers as in experiments 1A and 1B. Recognition rates across 3A and 3B in the post-test were again very high with an average of 93.0% (SD = 12.6) of faces recognized.

Identity was recognized at rates greater than chance for shape adaptation in Experiment 3A but not for color adaptation in Experiment 3B. One-sample t-tests against chance (50%) revealed that participants showed significant recognition of identity for experiment

3A (shape-only, t(21) = 2.81, *p* = .010) but not for experiment 3B (color-only, t(31) = -1.22, *p* = .230). An independent samples t-test revealed a significant difference in accuracy between Experiment 2A and 2B (t(52) = 2.99, *p* = .004). Means can be seen in Figure 6.

Figure 6 about here

Additional tests demonstrated that there was a weak effect of order such that accuracy was higher for earlier than later trials, though this effect was weaker in the color-only condition. One-sample t-tests against chance (0 = no correlation) revealed negative correlations between trial order and correct responses for both shape-only (mean r = -.10, SD = .20, t(21) = 2.30, *p* = .032) and color-only (mean r = -.08, SD = .24, t(31) = 1.79, *p* = .084) adapting images, though this was significant for the first and only tending towards significance for the second. Paired samples t-test revealed that participants were more accurate for the first 10 trials than the second 10 trials for the shape only images (1st 10: mean = 61.8%, SD = 15.9; 2nd 10: mean = 51.4%, SD = 13.9; t(21) = 2.47, *p* = .022) but not for the color only images (1st 10: mean = 51.3%, SD = 17.0; 2nd 10: mean = 44.4%, SD = 16.1; t(31) = 1.60, *p* = .119).

Discussion

Experiment 3A and 3B demonstrated that shape rather than color is mainly responsible for accurate identification after exposure to antifaces. This finding supports previous studies suggesting that objects may be encoded primarily in terms of their luminance-defined bounding edge structure (see Biederman, 1987) and research on faces demonstrating that observers are able to recognize familiar faces that have been hue-reversed at levels equivalent to normal-hued faces (Kemp et al., 1996). Faces generally have similar spectral properties and observed face color can change dramatically under different lighting conditions. As color may be unreliable across time, it is logical for the visual system not to rely on color as a diagnostic cue of identity. Some studies have found a role for color in identity recognition. Lee and Perrett (1997) found recognition accuracy for famous faces was greater when viewing color caricatured stimuli over veridical images and that the removal of color information also decreased accuracy of recognition. Color here then facilitates recognition when the face is already a recognizable identity and perhaps that is key to finding the color effect. In Lee and Perrett's study the faces already possess the correct celebrity shape and so the face is in the correct area of face space. Potentially only when the face has some other cues to identity may color aid recognition in ambiguous tasks. Another study has shown that the contribution of color cues becomes evident when shape cues are degraded and that the contribution of color may lie not in providing cues to identity but in aiding low-level processes (Yip and Sinha, 2002). Other studies, however, have highlighted the use of color in face recognition (Russell and Sinha, 2007). As face adaptation effects are proposed to reflect high-level processes then again it makes sense that color cues play a smaller role. That said, as the color images demonstrate the same order effect as the shape images then some similarities in the processes may be evident, though we note such effects are weak for color. We also note that color may be less important than shape cues if our faces are relatively homogeneous in their color cues but are more variable in shape.

Experiment 4

Here we examined the effects of using male versions of adapting stimuli. Previous studies have shown that face aftereffects can be simultaneously produced in opposite directions in male and female faces (Little et al., 2005), suggesting that faces are coded relative to a same-sex norm, rather than a sex-neutral norm. Sex specific effects of adaptation are also seen for identity for newly learned faces (Rhodes et al., 2011). Here we predicted that if adaptation to male faces has a limited influence on female face perception then adapting to anti-male faces that share the same properties of their anti-female face counterparts will not induce accurate perceptions of identity in the same way as the anti-female faces.

Participants

Participants in Experiment 4 (male adapting face) were 31 women and 17 men (mean age = 30.4, SD=10.2). Participants were selected as for Experiment 1A and 1B. The study was run over the internet and participants were recruited via a link from a research based website (www.alittlelab.com).

Stimuli

Stimuli were made in the same way as in Experiment 1A and 1B but here transformations were made against a composite androgynous celebrity image instead of the female composite and applied to the androgynous composite. This androgynous image was made by combining the female composite with a male equivalent made in the same way as the female image, from 20 individual composites of male celebrities each made from 10 different images. As before, this image was aligned and made symmetric prior to transform. Transforms were anti-100% in shape and color and each antiface contains the shape information apparent in the antifaces made through the female average. Moving through an androgynous shape means the antifaces also exhibit a level of apparent maleness akin and opposite to the femaleness of the individual celebrity. Example images can be seen in Figure 7.

Figure 7 about here

Procedure

The procedure was identical to that used in Experiment 1A.

Results

Recognition rates in Experiment 4 in the post-test were again very high with an average of 94.4% (SD = 14.3) of faces recognized.

Identity was not recognized at rates greater than chance when adaptaing across sex of face and the effect was significantly different from that seen for same-sex faces in Experiment 1A. A one-sample t-tests against chance (50%) revealed that participants did

not show correct recognition of identity in Experiment 4 (t(47) = -0.70, *p* = .489). We compared the accuracy scores with those from Experiment 1A and an independent samples t-test revealed that sex of adapting face (female versus male) significantly influenced accuracy (t(78) = 2.64, *p* = .010). Means can be seen in Figure 8.

Figure 8 about here

Additional tests demonstrated that there was no effect of order on accuracy when male faces were used as adapting stimuli. A one-sample t-test against chance (0 = no correlation) revealed no significant negative correlations between trial order and correct responses for male adapting antifaces (mean r = -.04, SD = .25, t(47) = 1.12, *p* = .269). Paired samples t-test revealed that participants were not more accurate for the first 10 trials than the second 10 trials for both the full-size images (1st 10: mean = 50.8%, SD = 18.8; 2nd 10: mean = 47.1%, SD = 15.3; t(47) = 1.03, *p* = .310).

Discussion

Experiment 4 demonstrated that transformations through an androgynous prototype, creating a male adapting face with the opposite shape and color characteristics, resulted in no accurate aftereffects for identity. Accuracy when using a male adapting face result was significantly worse than when using a female despite the fact that, bar the masculine appearance, shape and color information was the same between the two types of adapting antiface. Theoretically, like the female antifaces, adaptation should have resulted in perception of the correct identity following the correct trajectory through the average. Unlike the previous experiments, order had no influence on accuracy for male adapting images suggesting that identity aftereffects are specific to congruent-sex faces.

Sex congruency in facial aftereffects is consistent with previous studies that have demonstrated that adaptation effects can be separable by sex of face such that the perception of male and females can be pushed in opposite directions (Little et al., 2005;Bestelmeyer et al., 2008) and other studies showing identity aftereffects are greater for newly learned faces when same-sex faces are used in the pairs (Rhodes et al., 2011). Such sex-specific effects are consistent with the idea that, rather than comparing faces with a single prototypic face, observers have separable representations of male and female faces, which would allow such separable manipulation. Our findings here also support this notion as adaptation to antifaces does not result in accurate perception when using a non-congruently sexed face.

General Discussion

Our data add to a burgeoning literature showing that exposure to faces biases subsequent perceptions of novel faces (Leopold et al., 2001;Rhodes et al., 2001;Rhodes et al., 2003;Rhodes et al., 2004;Webster et al., 2004;Leopold et al., 2005;Little et al., 2005;Buckingham et al., 2006;Bestelmeyer et al., 2008;Little et al., 2008). The main finding across our studies was that adaptation to an antiface could make the same neutral average face appear to possess new identities specific to the identity of the antiface along

a particular trajectory in face space (Experiments 1-3). Such data strongly implies that the average face plays a special role in recognition of identity supporting prototype referencing models of face recognition (see also Leopold et al., 2001;Leopold et al., 2005;Rhodes and Jeffery, 2006). As noted earlier, this does not necessarily imply the prototype is explicitly mentally represented - the encoding of faces could involve contrastive neural mechanisms that reference the central tendency of the stimulus category (Rhodes and Jeffery, 2006). Such data also supports computer modeling studies suggesting prototype abstraction is a consequence of discrimination learning (Enquist and Arak, 1994;Johnstone, 1994). While the majority of the experiments presented here were web-based, our Experiment 2 demonstrates the equivalency of results from web- and laboratory-based testing for this type adaptation experiment.

We found that shape information was of greater importance than color in coding face identity (Experiment 3A and 3B). This is consistent with previous studies suggesting color is not crucial in recognizing objects (Biederman, 1987) or faces (Kemp et al., 1996) and that face color can change quickly. Of course, while adaptation here appears reliant on shape, possibly because the color cues alone do not allow the face to be disambiguated from other similarly colored identities which share the same color properties, alongside shape cues to identity color may help identify faces (Lee and Perrett, 1997). Indeed some studies do find that color plays an important role in face recognition (Russell and Sinha, 2007).

Previous studies have shown category contingent aftereffects whereby exposure can influence different categories in different ways simultaneously suggesting separable representation (Little et al., 2005;Bestelmeyer et al., 2008;Little et al., 2008). The results of Experiment 4 support the notion that humans have distinct representations for male and female faces as we were unable to influence the perception of female faces in the correct direction with adaptation to a male face with same antishape as an equivalent female antiface and are consistent with similar effects seen for newly learned faces (Rhodes et al., 2011). One study has demonstrated that adaptation to some aspects of visual appearance can cross from male to female faces (Jaquet and Rhodes, 2008). Though our data provide no support for the notion that aftereffects have any influence on perception across the category of sex, if adaptation can cross the category of sex, our results suggest across-sex effects are much weaker than for sex-contingent adaptation. It is commonly conceived that we recognize faces based on deviations from a single average representation within a single population of all faces encountered (Valentine, 1991), while our data and the studies cited above generally support the notion that perhaps we compare or relatively code faces against a specific prototype for each category.

One issue that was apparent across all of our Experiments (bar Experiment 4, which showed a null effect for adaptation to men’s faces) was the effect of order: adaptation to antifaces resulted in accurate perception of identity most strongly for those faces seen early in the trials and less so for faces seen in later trials. One trivial explanation is simple trial fatigue such that individuals paid less attention to later trials. This is an unsatisfactory explanation given the trials were short, as was the overall test,

and involved only the choice out of two alternatives. The task itself was also a relatively interesting one. Another explanation, while speculative, is that the order effect might reflect that the malleability we see in face perception after exposure is a finite resource so that repeated exposure to new faces results in weaker influences of later faces. It is possible, for example, that adapting to particular face shapes prevents later adaptation to new faces as the initial adaptation effects persist beyond the probe trials and this carry-over adaptation interferes with later adaptation. Further research is needed to investigate this potentially important issue for our understanding of flexible face processing.

Studies of adaptation also have potential implications for the neural representation of faces. Neural responses to identity are sensitive to differences across identity rather than physically equivalent within-identity changes (Rotshtein et al., 2005) and so the average face may then play an important role in helping categorize identity in terms of defining whether a face is more or less similar to another by defining a trajectory in face space. Indeed, neuroimaging and single-cell recording studies have also supported the notion of a prototype-referenced model of face coding (Giese and Leopold, 2005;Loffler et al., 2005). In future experiments measurement of the neural responses to adaptation to shape and color independently, and the effects of repeated adaptation to different faces will prove enlightening in further understanding the representation of faces. We note that although the results point to the locus of adaptation being mechanisms for coding high level aspects of faces, further work is also needed to fully rule out lower level explanations/loci. While our data demonstrate the effects of adaptation on identity recognition, adaptation effects are seen for many aspects of face perception including sex and emotion classification as well as overall perceptions of normality. Alongside such effects, adaptation is also clearly relevant to low level visual properties, such as blur, and there are commonalities in the properties of low and proposed higher level adaptation effects (Webster and MacLeod, 2011). Adaptation effects may then be a general feature of face processing applicable at several different levels of mental representation. An additional issue for prototypic mental representations of faces is how such representations are constructed from different views and image sizes. The distinction between low and high level adaptation effects and how the human visual system deals with variation in face stimuli in forming mental representations remain interesting avenues for future research.

Conflict of interest:

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.

Acknowledgments:

Anthony Little is supported by a Royal Society University Research Fellowship. We thank D.I. Perrett and B.P. Tiddeman for the use of their software.

References

Anderson, N.D., and Wilson, H.R. (2005). The nature of synthetic face adaptation. *Vision Research* 45**,** 1815-1828.

Ashby, E.G., and Maddox, W.T. (2005). Human category learning. *Annual Review of Psychology* 56**,** 149-178.

Benson, P.J., and Perrett, D.I. (1993). Extracting prototypical facial images from exemplars. *Perception* 22**,** 257-262.

Benson, P.J., and Perrett, D.I. (1994). Visual Processing of Facial Distinctiveness. *Perception* 23**,** 75-93.

Bestelmeyer, P.E.G., Jones, B.C., Debruine, L.M., Little, A.C., Perrett, D.I., Schneider, A., Welling, L.L.M., and Conway, C.A. (2008). Sex-contingent face aftereffects depend on perceptual category rather than structural encoding. *Cognition* 107**,** 353-365.

Biederman, I. (1987). Recognition-by-Components - a Theory of Human Image Understanding. *Psychological Review* 94**,** 115-147.

Buckingham, G., Debruine, L.M., Little, A.C., Welling, L.L.M., Conway, C.A., Tiddeman, B.P., and Jones, B.C. (2006). Visual adaptation to masculine and feminine faces influences generalized preferences and perceptions of trustworthiness. *Evolution and Human Behavior* 27**,** 381-389.

Enquist, M., and Arak, A. (1994). Symmetry, Beauty and evolution. *Nature* 372**,** 169-172.

Giese, M.A., and Leopold, D.A. (2005). Physiologically inspired neural model for the encoding of face spaces. *Neurocomputing* 65**,** 93-101.

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

Johnstone, R.A. (1994). Female preference for symmetrical males as a by-product of selection for mate recognition. *Nature* 372**,** 172-175.

Jones, B.C., Debruine, L.M., and Little, A.C. (2008). Adaptation reinforces preferences for correlates of attractive facial cues. *Visual Cognition* 16**,** 849-858.

Jones, B.C., Feinberg, D.R., Bestelmeyer, P.E.G., Debruine, L.M., and Little, A.C. (2010). Adaptation to different mouth shapes influences visual perception of ambiguous lip speech. *Psychon Bull Rev* 17**,** 522-528.

Kemp, R., Pike, G., White, P., and Musselman, A. (1996). Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception* 25**,** 37-52.

Lee, K.J., and Perrett, D. (1997). Presentation-time measures of the effects of manipulations in colour space on discrimination of famous faces. *Perception* 26**,** 733-752.

Leopold, D.A., O'toole, A.J., Vetter, T., and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience* 4**,** 89-94.

Leopold, D.A., Rhodes, G., Muller, K.M., and Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society B-Biological Sciences* 272**,** 897-904.

Little, A.C., Burt, D.M., Penton-Voak, I.S., and Perrett, D.I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society B-Biological Sciences* 268**,** 39-44.

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. *Proceedings of the Royal Society B-Biological Sciences* 272**,** 2283-2287.

Little, A.C., Debruine, L.M., Jones, B.C., and Waitt, C. (2008). Category contingent aftereffects for faces of different races, ages and species. *Cognition* 106**,** 1537-1547.

Loffler, G., Yourganov, G., Wilkinson, F., and Wilson, H.R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience* 8**,** 1386-1390.

Nosofsky, R.A., and Zaki, S.R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology-Learning Memory and Cognition* 28**,** 924-940.

Perrett, D.I., Lee, K.J., Penton-Voak, I.S., Rowland, D.R., Yoshikawa, S., Burt, D.M., Henzi, S.P., Castles, D.L., and Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394**,** 884-887.

Posner, M.I., and Keele, S.W. (1968). On Genesis of Abstract Ideas. *Journal of Experimental Psychology* 77**,** 353-363.

Rhodes, G., Brennan, S., and Carey, S. (1987). Identification and ratings of caricatures - implications for mental representations of faces. *Cognitive Psychology* 19**,** 473-497.

Rhodes, G., Halberstadt, J., and Brajkovich, G. (2001). Generalization of mere exposure effects in social stimuli. *Social Cognition* 19**,** 57-70.

Rhodes, G., Jaquet, E., Jeffery, L., Evangelista, E., Keane, J., and Calder, A.J. (2011). Sex-specific norms code face identity. *Journal of Vision* 11.

Rhodes, G., and Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research* 46**,** 2977-2987.

Rhodes, G., Jeffery, L., Watson, T.L., Clifford, C.W.G., and Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science* 14**,** 558-566.

Rhodes, G., Jeffery, L., Watson, T.L., Jaquet, E., Winkler, C., and Clifford, C.W.G. (2004). Orientation-contingent face aftereffects and implications for face-coding mechanisms. *Current Biology* 14**,** 2119-2123.

Ross, D.A., Hancock, P.J.B., and Lewis, M.B. (In Press). Changing faces: Direction is important. *Visual Cognition*.

Rotshtein, P., Henson, R.N.A., Treves, A., Driver, J., and Dolan, R.J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience* 8**,** 107-113.

Rowland, D.A., and Perrett, D.I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications* 15**,** 70-76.

Russell, R., and Sinha, P. (2007). Real-world face recognition: The importance of surface reflectance properties. *Perception* 36**,** 1368-1374.

Smith, J.D., and Minda, J.P. (2002). Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. *Journal of Experimental Psychology-Learning Memory and Cognition* 28**,** 800-811.

Tiddeman, B.P., Burt, D.M., and Perrett, D.I. (2001). Prototyping and transforming facial texture for perception research. *IEEE Computer Graphics and Applications* 21**,** 42-50.

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

Webster, M.A., Kaping, D., Mizokami, Y., and Duhamel, P. (2004). Adaptation to natural facial categories. *Nature* 428**,** 557-561.

Webster, M.A., and Macleod, D.I.A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B-Biological Sciences* 366**,** 1702-1725.

Yip, A.W., and Sinha, P. (2002). Contribution of color to face recognition. *Perception* 31**,** 995-1003.

Figure 1: Left: Examples of composite celebrity images (Jennifer Aniston and Angelina Jolie), anti-100% faces used as adapting stimuli and the average female celebrity composite used as a probe face. Antifaces are created by transforming the original face through the average face in both shape and color. Right: Schematic example of face space showing antifaces 'on the other side of the mean'. Arrows designate direction of transform.

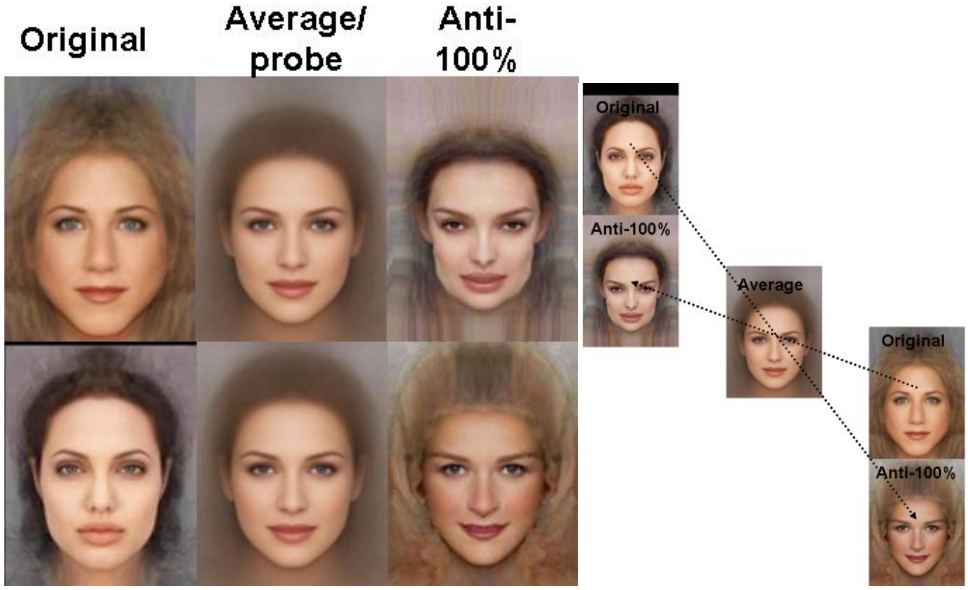

Figure 2: Trial structure for full size images (top) and 70% size images (bottom) as probe faces. The antiface for Jennifer Aniston is used here.

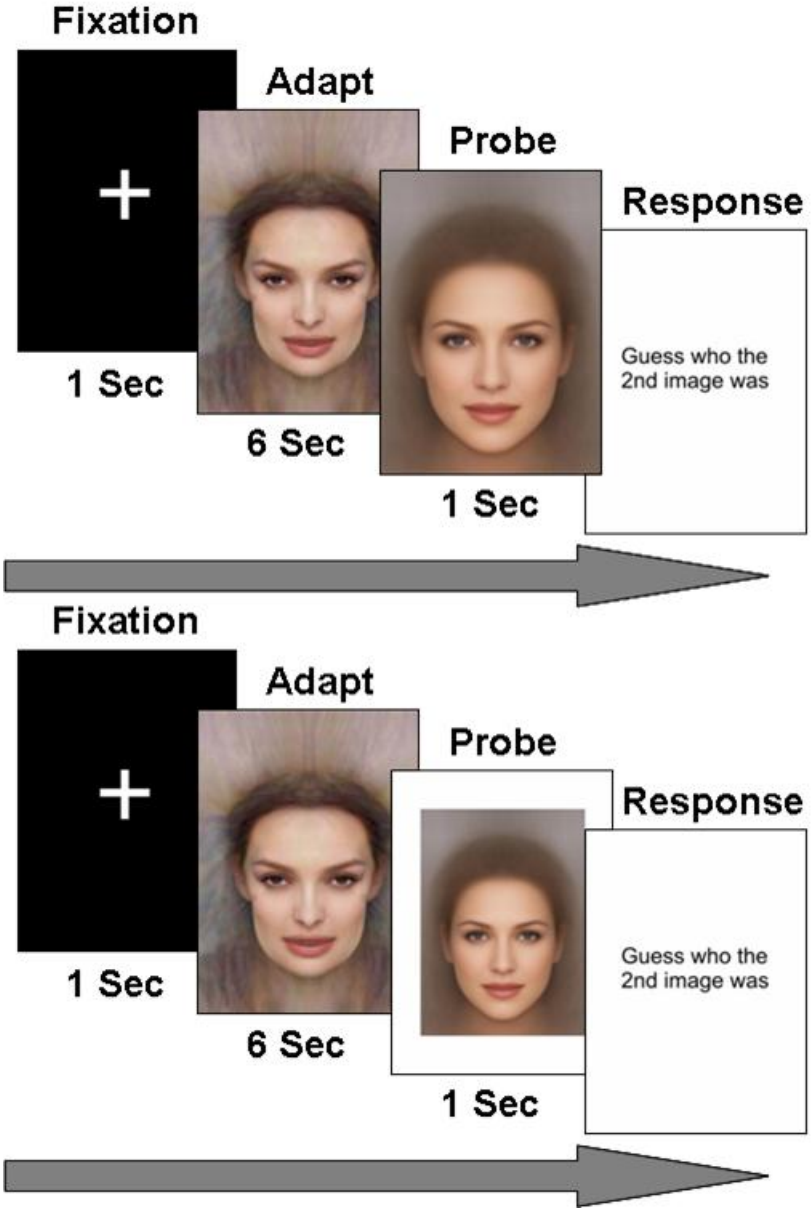

Figure 3: Correct responses (as %, +/- 1SE of mean) for Experiment 1A (same-sized probe) and 1B (70% sized probe).

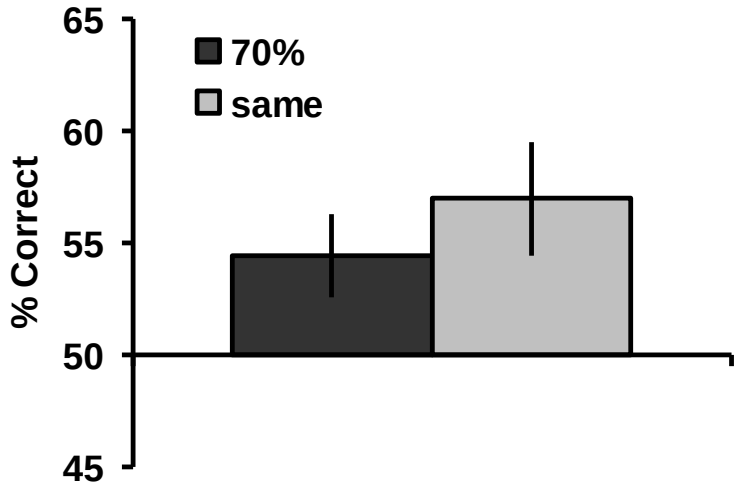

Figure 4: Correct responses (as %, +/- 1SE of mean) for trials where the test was conducted in the laboratory (Experiment 2) or over the web (Experiment 1A).

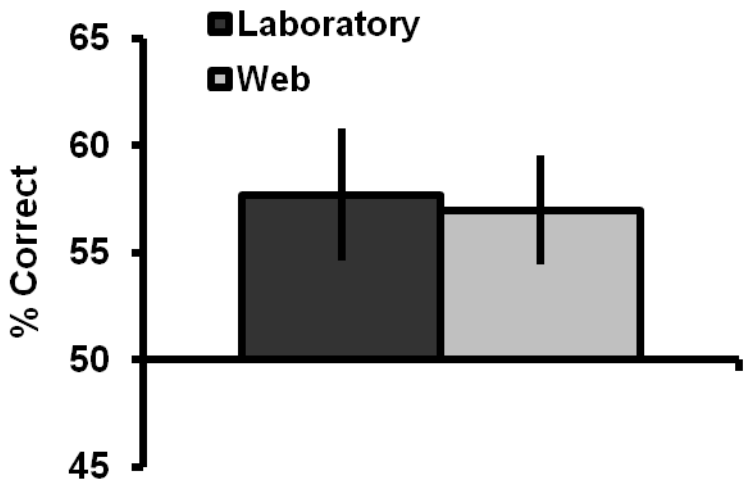

Figure 5: Examples of composite celebrity images (Jennifer Aniston and Angelina Jolie), anti-100% faces used as adapting stimuli and the average female celebrity composite used as a probe face. Antifaces are created by transforming the original face through the average face in either shape or color.

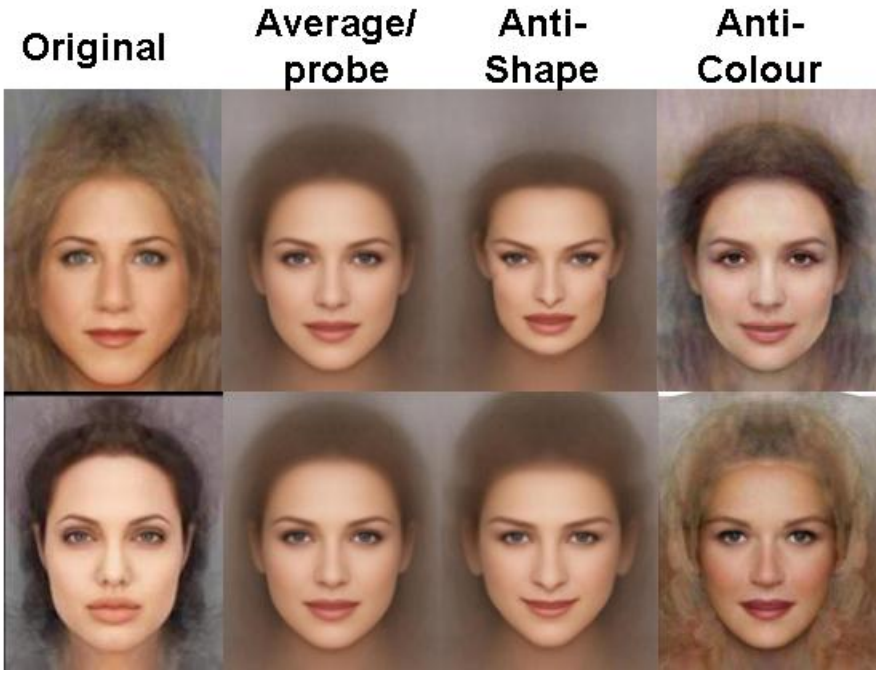

Figure 6: Correct responses (as %, +/- 1SE of mean) for trials where the adapting image was transformed in shape only (Experiment 2A) or the adapting image was transformed in color only (Experiment 2B).

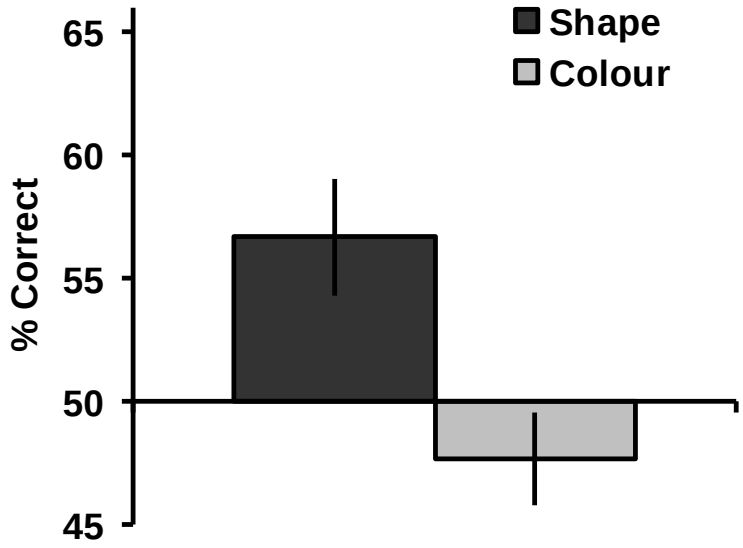

Figure 7: Examples of composite celebrity images (Jennifer Aniston and Angelina Jolie), the androgynous celebrity composite used for transforming and as a probe face, and male versions of the antifaces used as adapting stimuli. Antifaces here are created by transforming the original face through the androgynous average face - 100%.

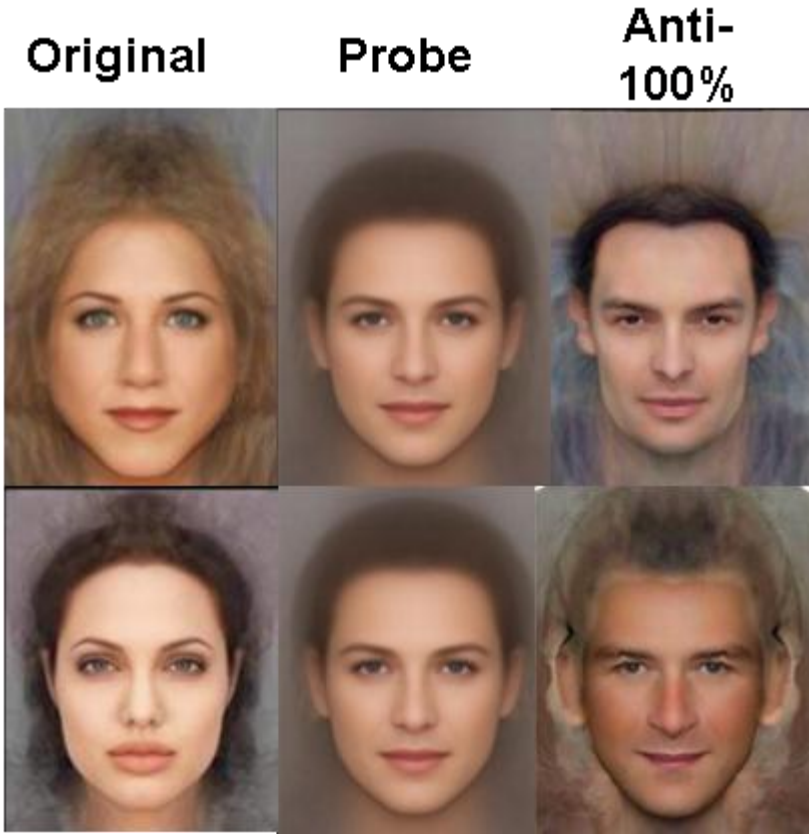
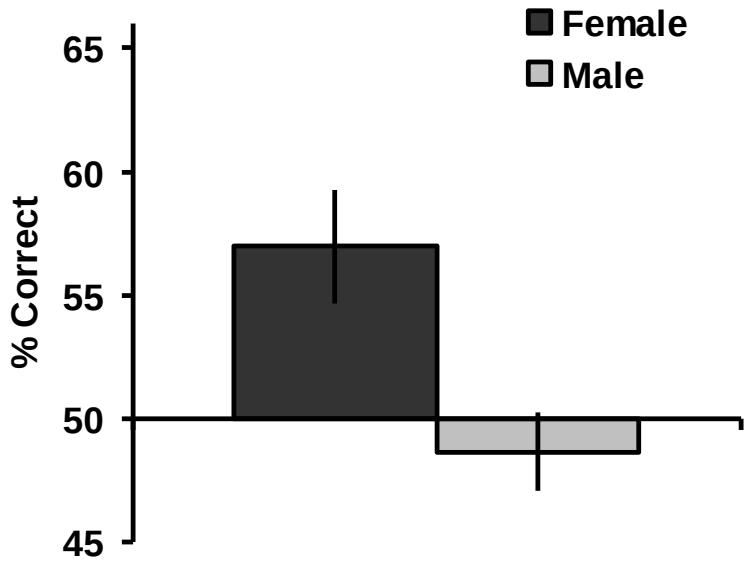
Figure 8: Correct responses (as %, +/- 1SE of mean) where adapting image was female and the probe the average female (data from Experiment 1A) and where the adapting image was male and the probe image was androgynous (data from Experiment 4).