Perception, 2011, volume 40, pages 317 – 322

Adaptation may cause some of the face caricature effect

Peter J B Hancock, Anthony C Little
Psychology, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, Scotland, UK;
e-mail: pjbh1@stir.ac.uk
Received 2 November 2010, in revised form 22 February 2011

Abstract. One of the ways to demonstrate a caricature preference is to ask participants to adjust a face image over a range from anti-caricature to caricature until it shows the best likeness to a specific individual. Since facial adaptation, whereby exposure to a face influences subsequent perception of faces, is rapid, it is possible that adaptation promotes the selection of a caricatured image. We tested whether giving participants a reference average face image, to counteract any adaptation, would reduce the degree of caricature selected for famous faces. Results confirmed a significant decrease but, even without an average, participants chose an anti-caricatured image. These data suggest a role for adaptation in generating caricature preferences while also suggesting such preferences are not inevitable.

1 Introduction

Caricaturing a face, by emphasising what is distinctive about it, has long been a tool of cartoonists. A face can be caricatured automatically, by shifting it away from an age/sex/race appropriate average face (Brennan 1985). Caricature effects can be studied via reaction time to recognition or by selection of the best likeness from a range of images. Early work used line-drawings of faces and therefore changes of shape only; Rhodes et al (1987) found that 50% caricatures (where the distance of each point from the average is increased by half) were recognised faster than veridical images. Benson and Perrett (1994) asked users to select the best likeness of a celebrity and found that, on average, a 42% caricature was chosen. This varied with the distinctiveness of the faces, with more distinctive faces being given less caricature. Benson and Perrett also reported an increase in speed and accuracy of naming the caricatures chosen as best likenesses.

However, line drawings are not very good representations of faces (Bruce et al 1992). Advances in computer graphics allowed caricatures to be generated from photographic images and with these more realistic images, evidence for a caricature preference is more equivocal. Benson and Perrett (1991) found that users selected on average only a 4.4% caricature of photographs of famous faces, while Benson and Perrett (1994) cite unpublished results with personally familiar faces where the preference was zero (ie the veridical image was chosen). There is some evidence for a recognition advantage for caricatured photographs, but only at very short exposure durations (33 ms—Lee and Perrett 1997). Lee and Perrett also asked users to select the best likeness of celebrities, and found an average of -8.2% (ie a shift towards the average image, an anti-caricature preference) for shape variations, but $+14.5\%$ for variation in the image (colour) domain. Image caricatures effectively increase the contrast of facial features; thus, if someone has darker than average eyebrows, the caricature will emphasise this. Kaufmann and Schweinberger (2008) asked their participants to judge the best likeness of shape variations of both celebrity and personally known faces. For both, they found no preference between veridical and anti-caricatures, but a reduced likeness rating for caricatures. Recently, Allen et al (2009) asked participants to adjust shape-only caricatures of themselves and of close friends and again found an average preference for a negative setting, -13.9%.

Such results are of interest partly for the bearing they have on how we may represent faces in memory. A preference for caricatures might imply that we store caricatured versions. Lewis and Johnston (1999) showed how this might arise as a natural consequence of the distribution of faces with a simple learning model. However, there are two confounding effects that have received little consideration. The first is that caricaturing alters not only changes in identity, but also all other deviations from the reference average face. Thus, if a particular face is rotated slightly to one side, a shape caricature will exaggerate this, producing distortions in the face. If the lighting differs, then an image caricature will emphasise this, along with any identity changes. It is particularly difficult to obtain celebrity images with a neutral expression, pose, and lighting. Since caricature emphasises all deviations, users asked to select the best likeness may select a lower degree of caricature in order to reduce such artifacts. Allen et al (2009) were able to avoid many of these problems by taking photographs of their targets under controlled conditions and they found preference for a negative caricature, but they tested only shape, not image, variation. Here we use average images of celebrities, produced by blending ten different images of each, which help to remove variations in lighting, etc, and therefore produce relatively unconfounded caricatures (see figure 1 for an example).

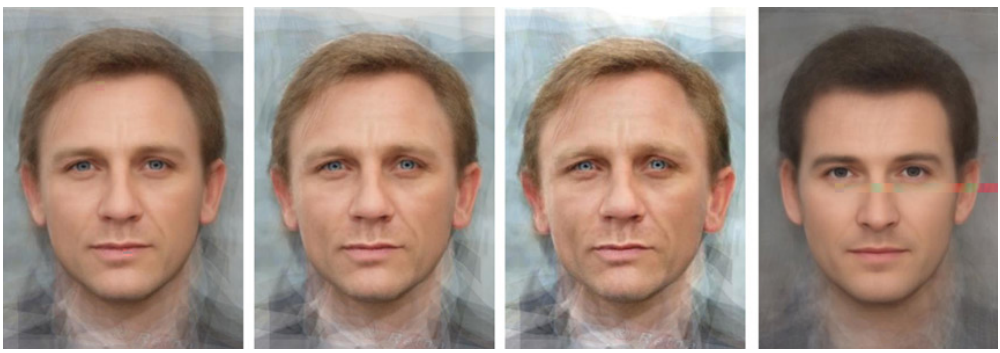

Figure 1. [In colour online, see http://dx.doi.org/10.1068/p6865] −30%, veridical average, and +30% caricatures of Daniel Craig, and the overall average male used as a reference

The second unconsidered issue is the possibility of adaptation having an effect on the selection. Observation of a given face rapidly produces a shift in perception; thus, looking at a female face makes an androgynous face look male, and vice-versa (Webster et al 2004). Leopold et al (2001) showed that the effect extends to the perception of identity. They created 'anti-faces' by reflecting a given face image through the mean; thus, someone with a bigger than average nose would end up with a small one; dark eyebrows become light, etc. They then tested participants' sensitivity to the identity of the original faces by testing their recognition of a continuum between the mean face and the veridical. After only 5 s adapting to the anti-face, sensitivity to the original identity was increased, such that recognition occurred closer to the mean. Now consider the implications of this for an experiment where participants are asked to adjust a caricature for the best likeness. Since the process will take several seconds, we can expect adaptation effects to alter judgments, in the direction of weakening the perceived identity (ie the opposite of adapting to an anti-face). The consequence is that users ought to choose a caricature, simply because they have been looking at the face for several seconds.

We therefore have two potentially opposing confounds: minimising image distortions might cause users to select more average images, while adaptation effects might cause them to select stronger caricatures. We address distortions by using average celebrity images; our approach to reducing the effects of adaptation is to present the overall average

face used to create the caricature alongside the celebrity image and ask participants to look between them, especially before making a final selection. While this cannot be expected to remove adaptation effects completely, it should reduce them, so we would therefore predict that users will select a less caricatured version of each face when the average reference image is present.

2 Method

2.1 *Participants*

Thirty-five students from the University of Stirling, fourteen male, mean age 20.3 years, took part voluntarily.

2.2 *Materials*

10 male and 10 female celebrities were selected as being well known to the likely participants. 10 frontal, non-smiling images of each were found from the internet, and each 'marked up' by locating 179 points around the outline of the face and the internal features, using PsychoMorph (Tiddeman et al 2001). These were then combined to give an average image for each celebrity. Overall average male and female celebrity images were computed from 29 such male averages and 48 females. PsychoMorph was then used to create a sequence of 11 images for each celebrity, in equal steps from 30% anti-caricature to 30% caricature, altering both shape and colour. Each image was 420×595 pixels; examples are shown in figure 1.

2.3 *Procedure*

Participants were tested individually in a cubicle. After explanation, they were presented with the first celebrity image. The program selected one of the 7 middle images of the sequence, ie somewhere in the range $\pm 24\%$ caricature to start from. This prevented always starting with the veridical image, while giving at least 2 more extreme images in either direction. Participants pressed the right and left arrow keys to change the image, which wrapped around when one of the end-points was reached. Therefore, pressing one key continuously would cycle through all 11 images, rather than sticking at the end, and the keys did not have a fixed meaning; either could both increase and decrease caricature. The participant was asked to choose the image they thought "looked most like the person depicted" by then pressing the space bar, or else to press 'q' if he/she did not know who it was. 10 of the celebrity images, half of each sex, were chosen at random from the set for this first stage of the experiment. There was then an instruction screen, explaining that there would now be an average image present to the left of the celebrity image and that participants were to refer to it for comparison, especially before making their final decision. They were then presented with the other 10 celebrity images as before, except for the presence of a sex-appropriate overall average image, of the same size, on the left of the screen. On completion, the theory underlying the experiment was explained to the participant. The whole study took around 5 minutes to complete.

3 Results

In total, across all participants, 700 celebrity images were shown, and 'q' was pressed 51 times, to indicate the celebrity was unknown. However, one participant pressed 'q' every time and had evidently misunderstood the instructions. Across the remaining thirty-four participants, therefore, they recognised 95.4% of the celebrities shown.

For those that were recognised, the average level of caricature with no average reference image was -9.5% (SD $= 9.5\%$); with a reference average, -14.4% (SD $= 10.1\%$). A paired t-test showed this difference to be significant when comparing mean scores by participant ($t_{33} = 2.88$, $p = 0.007$; effect size $d = 0.50$). The effect was more consistent

across items ($t_{19} = 3.51$, $p = 0.002$; $d = 1.16$). Adding the average reference image reduced the amount of caricature selected or, rather, increased the amount of anti-caricature.

The selections without an average were always made before those with an average, in order to prevent participants thinking about the average image during the no-average condition. A possible explanation for our results is that there is simply a systematic reduction in caricature level chosen across the 20 items. Figure 2 shows the average caricature level chosen arranged by serial order. We note that: for each serial position, the average chosen with a reference image is lower (more negative) than without; and the very first image shown is the least anti-caricatured. There appears to be little evidence overall for a serial-order effect. As the first image appears unusual, this item was removed and the analysis rerun. The difference remains significant ($t_{33} = 2.61$, $p = 0.013$).

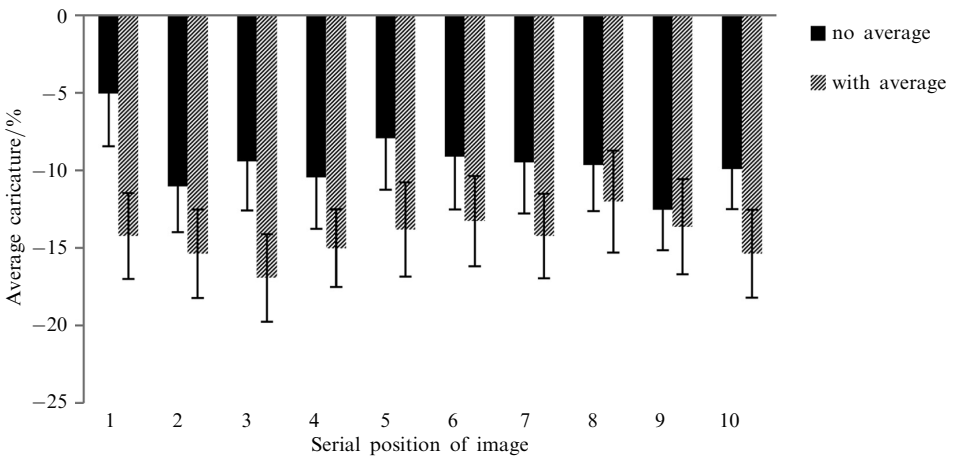

Figure 2. The average level of caricature chosen by serial order of image shown, without and with an average reference image. The serial position shown is within block, so the first image shown with an average is actually the eleventh image overall. Error bars are ±1 SE.

4 Discussion

As predicted, giving participants a reference average image caused them to choose less caricatured images as being most like the celebrity depicted. Rather against our initial expectation, but in line with the findings of Allen et al (2009), the overall preference was for anti-caricatured images, but more so with the reference image. The difference that we find between conditions is 4.9%, which is interestingly close to the original estimate of +4.4% for the preferred shape caricature found by Benson and Perrett (1991). The closeness of this match is probably coincidental, given the differences between the stimuli used (photographs changed by shape only, compared with ours changed by shape and colour), but it seems at least possible that their reported effect is caused by adaptation. Note, however, that adaptation clearly cannot be responsible for the improvement in recognition of caricatured photographs at 33 ms presentation (Lee and Perrett 1997).

Our results are consistently negative, however: people believe mild anti-caricatures to be most like the celebrities. In this we echo the findings of Allen et al (2009) and Kaufmann and Schweinberger (2008), who altered shape only, and Lee and Perrett (1997) when they altered shape, but not when they caricatured in the image domain, for which they found a preference for positive caricatures. One possibility is that shape effects dominate image changes. Another possibility is the presence of residual averaging artifacts in our images, owing to such things as minor inaccuracies in labelling

the key points. When starting this work, we thought a logical follow-up study would be to use faces personally familiar to the participants. These could be collected under carefully controlled conditions, allowing caricatures to be generated without emphasising photographic differences. However, this study has now been done by Allen et al (2009) who found a preference for an anti-caricature of around −14%, which is slightly stronger than our average of −9.5% without a reference image. The two studies complement each other, producing similar results with different types of images and different manipulations.

So why the negative preference? Allen et al make two suggestions. One is that we may be more able to discriminate small changes on the positive caricature side of veridical better than small changes on the negative side. The effect of this would be to make viewers tend to the negative. However, Dakin and Omigie (2009) explicitly tested the detectability of changes along the identity continuum, and found evidence for enhanced sensitivity on the average side of veridical, which would predict preference for a positive caricature.

Allen et al's second suggestion is that their result might be caused by an idealisation by their participants of the facial attractiveness of their own and their friends' faces, resulting in them choosing a more average, and therefore possibly more attractive variant, since average faces tend to be attractive (Langlois and Roggman 1990). Our results might argue against that interpretation, since we are using celebrities who are distinctively attractive. It is not the case that moving every face towards average will make it more attractive: the average of attractive faces is more attractive than the average of all faces (Perrett et al 1998), and moving in the direction of the more attractive faces increases attractiveness. An over-idealisation of the attractiveness of the celebrity faces might be expected to cause a preference towards a caricature, to emphasise what it is about their faces that is attractive. To test this directly, we ran a simple rating study, asking twenty-one women to rate how attractive they found our 20 celebrity pictures, which were either veridical, or ±12% shape caricatures, counterbalanced across raters. On a scale from 1–7, with 7 attractive, the anti-caricatures averaged 4.54, the veridical images 4.34 and the caricatures 4.1. The anti-caricatures were significantly more attractive than the caricatures ($t_{20} = 2.49$, $p = 0.02$). Whether this is because they are more average, or because they reduce distortions caused by averaging, we cannot say, but the finding is consistent with Allen et al's interpretation.

A further possibility, suggested by an anonymous reviewer, is that the anti-caricatures look younger and that, in our memories, celebrities are also younger than they really are. This might cause viewers to select an anti-caricature in order to reduce the apparent age. To test the referee's intuition that the anti-caricatures look younger, we asked twelve colleagues to estimate the age of the faces depicted, counterbalanced as for the attractiveness rating. The mean age estimate for the anti-caricatures was 31.8, compared with 34.1 for the veridical images and 33.7 for the caricatured. By items, the anti-caricatures are indeed rated as significantly younger than both the other sets ($t_{19} = 3.13$, $p = 0.006$, and $t_{19} = 2.44$, $p = 0.025$, respectively).

In conclusion, we have provided evidence suggesting that adaptation may cause viewers to select more caricatured (or, at least, less anti-caricatured) images as being veridical. However, it is apparent that there are many other possible factors affecting the choice of image, including imaging artifacts, detectability of changes, and other preferences of the viewer, such as perceived attractiveness and age. Some of these issues may be addressed by a move to 3-D face imagery, though there remain significant issues with averaging these. Overall, evidence for a 'caricature effect' for photographic images, beyond the small increase in recognition at 33 ms presentation reported by Lee and Perrett (1997), seems weakened by data from both personally known and famous faces, suggesting the opposite effect: anti-caricatures are chosen more often than caricatures.

Acknowledgments. Primary data were collected by Sam Baker, Olivia Childs, Amanda Farr, Storm Rae, and Maggie Sullivan.

References

Allen H, Brady N, Tredoux C, 2009 "Perception of 'best likeness' to highly familiar faces of self and friend" *Perception* **38** 1821 – 1830

Benson P J, Perrett D I, 1991 "Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images" *European Journal of Cognitive Psychology* **3** 105 – 135

Benson P J, Perrett D I, 1994 "Visual processing of facial distinctiveness" *Perception* **23** 75 – 93

Brennan S, 1985 "Caricature generator: The dynamic exaggeration of faces by computer + illustrated works" *Leonardo* **18** 170 – 178

Bruce V, Hanna E, Dench N, Healey P, 1992 "The importance of 'mass' in line drawings of faces" *Applied Cognitive Psychology* **6** 619 – 628

Dakin S C, Omigie E, 2009 "Psychophysical evidence for a non-linear representation of facial identity" *Vision Research* **49** 2285 – 2296

Kaufmann J M, Schweinberger S R, 2008 "Distortions in the brain? ERP effects of caricaturing familiar and unfamiliar faces" *Brain Research* **1228** 177 – 188

Langlois J H, Roggman L A, 1990 "Attractive faces are only average" *Psychological Science* **1** 115 – 121

Lee K J, Perrett D I, 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, Blanz V, 2001 "Prototype-referenced shape encoding revealed by high-level aftereffects" *Nature Neuroscience* **4** 89 – 94

Lewis M B, Johnston R A, 1999 "A unified account of the effects of caricaturing faces" *Visual Cognition* **6** 1 – 42

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

Rhodes G, Brennan S, Carey S, 1987 "Identification and ratings of caricatures: implications for mental representations of faces" *Cognitive Psychology* **19** 473 – 497

Tiddeman B, Burt D M, Perrett D I, 2001 "Prototyping and transforming facial textures for perception research" *IEEE Computer Graphics and Applications* **21** 42 – 50

Webster M, Kapling D, Mizokami Y, Duhamel P, 2004 "Adaptation to natural facial categories" *Nature* **428** 557 – 560

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION
VOLUME 40 2011

www.perceptionweb.com