Brief Report

When do Caricatures Look Good?

Gillian Rhodes
Department of Psychology, University of Otago

For line drawings of colleagues' faces more extreme caricatures look good (Rhodes, Brennan & Carey, 1987) than for photographic caricatures of famous faces (Benson & Perrett, 1991). The experiment reported here used line drawings of famous faces to see whether this discrepancy is due to the different kinds of faces used or to the different picture media used. The mean caricature level of the preferred picture was +5.5%, which is close to the +4.4% reported by Benson and Perrett, but much lower than the +16% reported by Rhodes, et al. Apparently more extreme caricatures look good for colleagues' faces than for famous faces, which is consistent with the idea that highly familiar faces (e.g., colleagues' faces) are "caricatured" in memory. Recognizability judgments were also obtained for pictures of birds — the undistorted drawings were preferred.

During Watergate, as Nixon's popularity plummeted, the length of his nose and the size of his jowls in published caricatures increased to impossible proportions. Yet the caricatures remained instantly recognizable. Sometimes caricatures may even be "superportraits" with the paradoxical quality of being more like the face than the face itself. Evidence of this "superportrait" idea comes from studies using computer-generated caricatures (Rhodes, Brennan & Carey, 1987), in which line drawing caricatures of colleagues (fellow graduate students and staff) were recognized more quickly than undistorted images.

Rhodes et al., speculated that this caricature advantage might occur because distinctive information about a familiar face becomes exaggerated or "caricatured" in long term memory, so that a caricature, which similarly exaggerates the distinctive information, would access the memory representation more quickly than an undistorted image. This speculation is consistent with several failures to find a caricature advantage for relatively unfamiliar faces or famous faces, both of which would be much less familiar than colleagues' faces. For relatively unfamiliar faces undistorted drawings are considered the best likeness (Rhodes et al., 1987) and undistorted drawings of previously presented faces are recognized more quickly than caricatures (Rhodes & Moody, 1990). When names are matched to famous faces, there is no advantage for caricatures over undistorted drawings (Benson & Perrett, 1991, Experiment 2).

If memory representations of familiar faces do become "caricatured" in long-term memory then one might expect caricatures to be accepted as good likenesses and possibly even as better likenesses than undistorted images. There is some support for this prediction. When shown 25%, 50% and 75% caricature of colleagues (together with the corresponding anticaricatures and the 0% drawings) Rhodes et al's subjects ranked the 25% and 50% caricatures equal with the undistorted (0%) images. The overall mean caricature level of the best likeness chosen was +16%, which was significantly higher than 0%.[1] Benson and Perret's (1991) subjects made "best likeness" judgments for photographic caricatures of famous faces (Experiment 1). Faced with (subsets of) 0%, 16% and 32% caricatures and anticaricatures their mean caricature level of the best likeness was +4.4%, which was also significantly higher than 0%. Slightly more of the subjects (just over .4) chose the 0% drawings as the best likenesses than chose the 16% drawings (just over .3) and very few subjects (about .1) chose the 32% drawings.

The overall pattern of results obtained in the two studies was similar. In both cases caricatures with a modest degree of exaggeration were considered to be as good likenesses (or almost as

I thank Craig Dowden, Wayne Edgeler, Richard Inder, Vicki Patton, Julie Richardson and Helen Stewart who assisted with data collection as part of a student research project. Requests for reprints should be sent to Dr Gillian Rhodes, Department of Psychology, Canterbury University, Private Bag 4800, Christchurch, New Zealand.

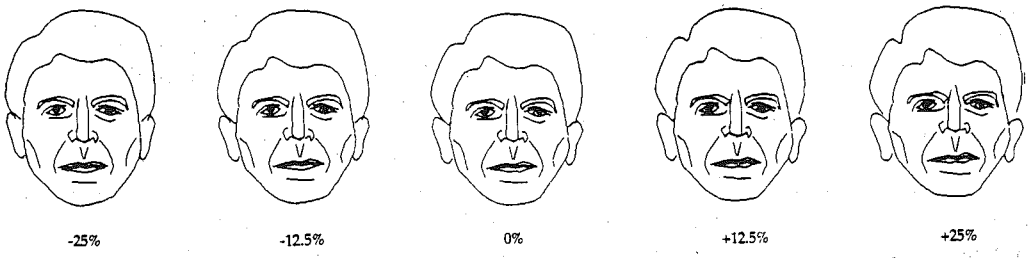

Figure 1. The full set of drawings for Ronald Reagan

good) as the undistorted faces. However, for the line drawings of colleagues' faces more extreme caricatures looked good (Rhodes, et al., 1987) than for the photographic caricatures of famous faces (Benson & Perrett, 1991). This difference could be due to the type of face (colleagues or famous faces). For example, colleagues' faces are seen more often and (usually) over a more extended time frame than are famous faces. The greater exposure might provide more opportunity for the memory representations of colleagues to become "caricatured", so that more extreme caricatures might be considered good likenesses for colleagues' faces than for famous faces. Alternatively, the difference in results could be due to the different type of pictures used. For example, the greater resemblance of photographic images to faces as normally seen might bias subjects towards choosing the veridical images a best likenesses.

The present study examined recognizability judgments for line drawings of famous faces in order to see whether the results would match the colleague/drawings data or the famous/photos data. A match with the former would indicate that the type of picture determines the degree of caricaturing that looks good and a match with the latter would indicate that the familiarity of the faces determines what looks good. The latter result would be consistent with the idea that memory representations change over time with faces becoming "caricatured" in memory as they become very familiar.

Judgments were also obtained for another set of stimuli, common birds. There is no caricature advantage for identifying birds, unless the subjects are bird experts (Rhodes & McLean, 1990), so there is no reason to expect the non-expert subjects in the present experiment to judge bird caricatures to be more recognizable than undistorted drawings. The bird data were collected to see whether differences in recognition performance for faces and birds would be reflected in differences in best likeness judgements.

Method

Subjects

Thirty-five unpaid volunteers participated in the experiment. All were undergraduate students or friends of students at the University of Otago.

Stimuli

Photographs of ten famous faces (see Figure 3 for names) and ten familiar birds (kiwi, kingfisher, pheasant, hawk, pelican, duck, puffin, swan, ostrich and penguin) were used in the experiment. Line drawing caricatures were created using Brennan's caricature generator (Brennan, 1982, 1985 — modified version was used for birds, Rhodes & McLean, 1990). The caricature generator produces a caricature in three steps. First, the image is digitized and a line drawing is created. The drawing is based on a fixed set of points (found by eye — 169 points for faces, 72 points for birds) that are joined automatically by the program using spline curves to produce a line drawing. The set of points was initially chosen to provide an economical, but recognizable, drawing. Second, the line drawing is compared to a norm (average) and third, all metric differences between the two are increased (or decreased for an anticaricature) by a specified percentage.

Five levels of caricature were created for each stimulus: –25%, –12.5%, 0%, 12.5% and 25%. Figure 1 shows the full set for Reagan. Negative values indicate anticaricatures in which distinctive information is reduced and positive values indicate caricatures in which distinctive information is exaggerated. The average of the three most typical male and female faces in a large set were used as the male and female norms, respectively (see Rhodes et al., 1987) and the average of the ten birds were used as the bird norm.

Procedure

At the beginning of the session subjects made familiarity ratings for each bird and face (based on names) on a 5-point scale (1 = "completely unfamiliar", 5 = "very familiar"). Then they saw the sets of five drawings

(randomly ordered from left to right) of each face (or bird). For each set they picked the "most recognizable" one, the next most recognizable one, and so on, until the five pictures were rank ordered. Pictures were blocked by type (face or bird) and half the subjects saw faces first and half saw birds first.

Results and Discussion

The mean rank was calculated for each type of picture (faces and birds) and each caricature level for each subject. Faces and birds rated 1 on familiarity (completely unfamiliar) were excluded from consideration. On average fewer than one face and bird were excluded per subject (0.7 out of 10 faces and 0.3 out of 10 birds). The mean familiarity level of the remaining faces and birds was 4.4 (SD = .6) and 4.3 (SD = .5), respectively. A three-way ANOVA was carried out with caricature level and picture type as within-subject factors and order (faces first or birds first) as a between subjects factor.

There was a significant main effect of caricature level, $F(4,132) = 35.07$, $p<.00001$, which was qualified by an interaction of caricature level with picture type, $F(4,132) = 26.76$, $p<.00001$ (see Figure 2). As predicted, for faces, but not for birds, some of the caricatures (12.5%) looked as good as the undistorted drawings. Planned comparisons (t-tests) confirmed that the mean ranks did not differ for 0% and 12.5% faces, but that 0% birds were preferred to both 12.5% and 25% bird caricatures ($p<.01$). For faces, caricatures were preferred over anticaricatures at both the 12.5% and 25% levels (both p's<.01, Tukey tests). In contrast, for birds the anticaricatures were preferred over the caricatures (ns. for 12.5%; $p<.01$ for 25%, Tukey tests). Therefore, the bird and face results differed, as has been found in recognition studies.

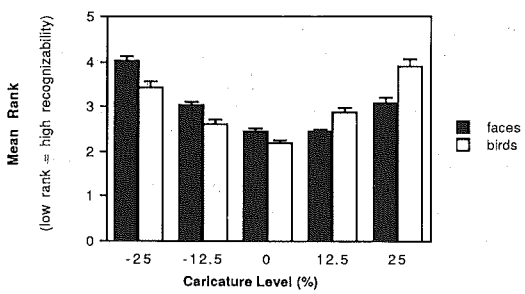

Figure 2. Recognizability choices: Mean rank as a function of caricature level and picture type. Standard error bars are shown.

There was also a significant 3-way interaction between picture type, caricature level and order, $F(4,132) = 3.35$, $p<.02$. The pattern displayed in Figure 2 held for both orders, but the difference in mean rank between birds and faces was greater at the 25% caricature and anti-caricature levels when birds were seen first than when faces were seen first. The interaction has no obvious theoretical significance and order did not alter the relative preferences for the different caricature levels described above. No other effects were significant.

The mean caricature level of the most preferred (i.e., lowest ranked) face was +5.6%, which was significantly greater than 0%, $t(34) = 3.58$, $p<.001$, 1-tailed. For birds the mean was –1.1, which did not differ significantly from 0%, $t<1$. A one-way ANOVA with picture type as a repeated measures factor showed that there was a significant effect of picture type, $F(1,34)=21.47$, $p<.0001$. The face result of +5.6% is almost identical to the +4.4% mean obtained for photographic caricatures of famous faces by Benson and Perrett (1991) and is much lower than the +16% mean obtained for line drawing caricatures of colleagues' faces by Rhodes, et al. (1987). Therefore the degree of exaggeration that looks good appears to be determined more by the type of face (colleague or famous) than the type of picture (line drawing or photograph). Less extreme caricatures of famous faces appear to look good than for colleagues' faces.

For eight of the ten faces in the experiment the mean caricature level of the "best likeness" (averaged across subjects) was greater than 0% (see Figure 3). For a couple of faces (Churchill and Reagan) the 12.5% caricatures appear to have been consistently preferred over the 0% drawings. However, for most of the faces the mean was about mid-way between 0% and 12.5%, indicating that these two levels were chosen as the most recognizable likeness about equally.

Taken together with the previous results (Benson & Perrett, 1991; Rhodes, et al., 1987) the present result for faces suggests that less extreme exaggeration looks good for famous faces than for faces of colleagues. The absence of a caricature preference for birds indicates that the face result obtained here is not an artifact of the way the stimuli were generated or presented.[2] The preference for more extreme caricatures of personally known than famous faces is consistent with the idea that highly familiar faces become "caricatured" in memory.

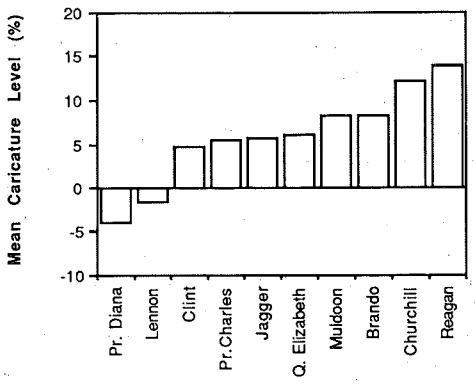

Figure 3. Mean caricature level of the "most recognizable" drawing (averaged across subjects) for each famous face.

Footnotes

1 Similar results were obtained when ranks rather than ratings where analyzed.

2 The studies with famous faces used a narrower range of caricature levels (0%, 16%, 32% in Benson & Perrett, 1990; 0%, 12.5%, 25% in the present study) than the study with colleagues' faces (0%, 25%, 50%, 75% in Rhodes et al., 1987). It is possible that this could influence the mean preferred level, but it would not affect the ranking data.

References

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.

Brennan, S. E. (1982). *Caricature generator.* MSc thesis, MIT, Cambridge MA.

Brennan, S. E. (1985). The caricature generator. *Leonardo, 18*, 170–178.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representation of faces. *Cognitive Psychology, 19*, 473–497.

Rhodes, G., & McLean, I. G. (1990). Distinctiveness and expertise effects with homogenous stimuli: Towards a model of configural coding. *Perception, 19*, 774–794.

Rhodes, G., & Moody, J. (1990). Memory representations of unfamiliar faces: Coding of distinctive information. *New Zealand Journal of Psychology, 19*, 70–78.