Caricaturing can improve facial expression recognition in low-resolution images and age-related macular degeneration

Jo Lane, Rachel A. Robbins, Emilie M. F. Rohan, Kate Crookes, Rohan W. Essex, Ted Maddess, Faran Sabeti, Jamie-Lee Mazlin, Jessica Irons, Tamara Gradden, Amy Dawel, Nick Barnes, Xuming He, Michael Smithson, and Elinor McKone

Journal of Vision, 3rd May 2019

SUPPLEMENTARY MATERIALS
SUPPLEMENT S1. Intensity rating experiment in young adults.
SUPPLEMENT S2. Details of blur for Experiment 1.
SUPPLEMENT S3. Additional procedure details for Experiment 1.
SUPPLEMENT S4. Additional results for Experiment 1: More severe effects of aging on "hard" than "easy" expression recognition.
SUPPLEMENT S5. Additional Experiment 2 details on patients and eyes: Inclusion criteria, excluded eyes, and more detailed vision testing. Includes Supplementary Table S2.
SUPPLEMENT S6. Additional procedure details for Experiment 2.
SUPPLEMENT S7. References for Supplementary Materials.
SUPPLEMENT S1. Intensity rating experiment in young adults.

Our split of the 82 Veridical faces into high, medium and low intensity of the expression was based on ratings obtained from young adult observers (a different group from those who participated in the main Experiment 1) tested individually.

Participants
Participants were 25 young adults reporting normal or corrected-to-normal vision (all Caucasian; 17 female, 8 male; age Mean = 21.8 years, SD = 4.1, range 18-38). Recruitment was via advertisement to the student community at the Australian National University. Participants received course credit or were paid $15 per hour; this covered the rating task and also another experiment not reported here (testing expression recognition in a bionic eye simulation). The time required per participant for the intensity rating experiment was 5-10 mins. Participants gave informed written consent after explanation of the nature and possible consequences of the study. The research methods adhered to the Declaration of Helsinki and were approved by the Australian National University Human Research Ethics Committee.

Stimuli and Procedure
Participants were shown the same 82 Veridical face stimuli used in Experiment 1, all in high resolution. The faces were shown one a time at screen centre until response, in a different random order for each participant. The task question was "How intense does this emotional expression look to you?" and the response scale (shown on the screen) was a 9-point scale running from 1 ("weak") to 9 ("strong").

Data Analysis and Results
For each individual face item, the intensity ratings from the 25 participants were averaged. The face items were then rank ordered from lowest to highest mean intensity rating, and divided into the lowest, middle and highest third. Properties of the resulting low, medium, and high intensity Veridical sets are described in main text Table 1.
SUPPLEMENT S2. Details of blur for Experiment 1.

The blur was implemented in the same way as in Irons et al. (2014). To quote details from that paper (pp. 9-10):

"We simulate the acuity–eccentricity relation by removing the frequency components higher than the cutoff frequency at an eccentricity e. We implement this by applying a uniform spatial blur across the image using a Gaussian kernel filter of size defined by the cutoff frequency. It is well known that the cutoff frequency $f$ (cpd) follows an inverse law with respect to the eccentricity (Anstis, 1974; Peli, Yang, & Goldstein, 1991; Rovamo, Virsu, & Näsänen, 1978). We represent this as $f = f_0/(1 + \beta e)$ and set the parameters as $\beta = 0.645$ and $f_0 = 30$ in this study. The resulting curve provides a close fit to the values presented in Marmor and Marmor (2010).

Achieving theoretically ideal frequency cutoff using image filtering is difficult to implement without introducing artifacts; we adopt the conventional Gaussian filters with a kernel width $\sigma_f$ as $(1/3)f$. This ensures that almost all the frequency components beyond $f$ will be removed. Given the value of $\sigma_f$ and the physical parameters used in the study, we convert the kernel width to pixel unit as follows. Let the width of screen be $w$ cm, the distance between viewer and screen be $d$ cm, and the horizontal resolution of the image be $r$ pixels. Based on Fourier transform theory, the Gaussian kernel width in the spatial domain is $\sigma_s = 1/(2\pi \sigma_f)$ degrees. Therefore, we can compute the filter kernel width as $\sigma_s = 3(1 + \beta e)/2\pi f_0 \times r/2\arctan(w/2d)$ pixel."
SUPPLEMENT S3. Additional procedure details for Experiment 1.

Concerning computer equipment and software:
- For young adults (tested in groups in a class setting), stimuli were displayed on Dell Optiplex 780 PC computers running Windows 8 with 24 inch (530x300 mm) Dell monitors with a screen resolution of 1920x1080. The experiment was run in Superlab 5.
- For older adults (tested individually), stimuli were displayed on a 27 inch imac running OSX, with a screen resolution of 2560x1400. The experiment was run in Superlab 4.5.
SUPPLEMENT S4. Additional results for Experiment 1: More severe effects of aging on "hard" than "easy" expression recognition.

The analysis of recognition accuracy for Veridical (i.e., uncaricatured) faces revealed, in addition to the main effects reported in the main article, three interactions.

Aging more severely affects "hard" than "easy" expression recognition. There was a significant two-way interaction between age group (young adults, older adults) and intensity of Veridical expressions (high, medium, low), $F(2,144)=5.825, MSE=59.793, p=.004$, partial $\eta^2=.075$. This interaction reflected (main text Figure 4) a stronger age-related reduction in recognition accuracy for low and medium intensity expressions than for high intensity expressions. This is of some potential theoretical interest, arguing that aging impairs "difficult" emotion recognition more strongly than "easy" emotion recognition. We note our results are for expression recognition from face-only information, without any context that improves emotion recognition (e.g., what a person displayed is looking at, what else is happening in a full scene, information about previous events leading up to the event; Aviezer, Trope, & Todorov, 2012). Our results raise questions such as whether, in everyday settings, older adults would also find it particularly hard to recognise subtle emotions and/or might need to rely more on context to do so.

Other interactions. The other significant interactions were of less interest. An intensity x blur interaction, $F(4,288)=12.455, MSE=47.676, p<.001$, partial $\eta^2=.147$, reflected the fact that, collapsed across age of observer, accuracy for medium and high intensity expressions was equal at high resolution, but medium was better than high at low resolutions. The direction of the effect (medium > high) was unexpected, and it was not replicated in Experiment 2 (with low resolution patient vision); we would wish to see the finding replicated with a different specific set of face stimuli before relying on it. Finally, a three-way age group x intensity x blur interaction, $F(4,288)=5.256, MSE=47.676, p<.001$, partial $\eta^2=.068$, appears to reflect simply ceiling and floor effects at the extremes of performance accuracy: that is, the aging effect disappears at the very top end of performance (i.e., for high resolution, high intensity faces), and also at the very bottom end of performance (i.e., for the most blurred, low intensity faces).
SUPPLEMENT S5. Additional Experiment 2 details on patients and eyes: Inclusion criteria, excluded eyes, and more detailed vision testing

Inclusion criteria and excluded eyes

Table S2 shows vision data for included and excluded eyes. As shown, we originally recruited 13 patients, however one patient (labelled Px in Table S2) failed to meet inclusion criteria for either eye, leaving the 12 patients whose demographics are described in the main text Experiment 2 Methods. Of the total of 26 eyes originally available: 1 eye was excluded due to not having AMD; 3 eyes were excluded due to having vision too poor to allow testing on the face recognition task (i.e., the patient reported they could not see the faces on the screen); and 3 eyes were excluded due to having vision that was too good, and thus having no need for image-enhancement technology.

These "too good" eyes all had AMD based on diagnosis of the retina, but demonstrated no relevant functional vision loss. One had acuity at or above normal vision levels, with BCVA = 6/4.8. Two had mild deficits in acuity (Patient Px’s right eye with BCVA = 6/7.5; and Patient Pd’s right eye with BCVA = 6/9.5) but performed at the 'maximum performance' level for Veridical faces in the expression recognition task: specifically, averaged across all 82 Veridical stimuli, expression recognition accuracy for these two eyes was 83% and 85% correct, which was above the mean from Experiment 1 of 82% in normal-vision young adults. The next-best-performing AMD eye scored 79% correct, which we considered far enough below maximum performance to be retained.

Detailed vision testing

Patients were given a complete vision assessment lasting 1.5 hrs. They gave informed written consent after explanation of the nature and possible consequences of this assessment. Research methods adhered to the Declaration of Helsinki and were approved by the Australian National University (ANU) and ACT Health Human Research Ethics Committees.

Table S2 includes LCVA and AREDS (2001) score. LCVA was measured using a retro-illuminated logMAR chart mounted on a stand conforming to the ETDRS format. AREDS stages are based on anatomy of the central 6mm of the retina (Stage 1 = Early AMD, small drusen; 2 = Early AMD, intermediate drusen; 3 = Early AMD, large drusen; 4 = covers active exudative, chorioidal neovascularisation for Wet AMD, and end-stage Dry AMD/sub-foveal geographic atrophy). For Stages 1-3 visual acuity is usually close to normal; for Stage 4, acuity can vary widely between normal and <6/60 (legally blind), e.g., depending on treatment (for Wet AMD).

Anterior segment of the eye was examined using slit-lamp biomicroscopy, instilling Oxybuprocaine Hydrochloride 0.4% eye drops to anesthetise the eyes to measure intraocular pressure using Goldmann applation tonometry and to measure central corneal thickness using a Pachmate (DGH Technology Inc., Exton, PA). Patients were tested on 10-2 frequency doubling technology (FDT) threshold using Humphrey Matrix (Carl Zeiss Meditec, Inc., Dublin, CA). After the visual field test both eyes were dilated with Tropicamide 1% and Phentylephrine 2.5% and the following tests were done: Optical Coherence Tomography (OCT) Spectralis (Heidelberg Engineering, Heidelberg, Germany) of the retina (posterior-pole) and the peripapillary retinal nerve fibre layer (pRNFL); scan to measure the thickness of the RNFL surrounding the optic nerve; fundus auto-fluorescence images were acquired, with fundus photography performed using a Canon CR-2 (Canon Inc. Medical Equipment Group, Tokyo, Japan) digital non-mydisriatic camera to get an image of the fovea, the macula and the optic nerve.
**SUPPLEMENTARY TABLE S2.** Participant vision information for both eyes.  

<table>
<thead>
<tr>
<th>Patient Code</th>
<th>Eye code (left or right)</th>
<th>Visual Acuity</th>
<th>Diagnosis</th>
<th>AREDS Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pa</td>
<td>E1 (R)</td>
<td>6/7.5</td>
<td>Early AMD</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>E5 (L)</td>
<td>6/9.5</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pb</td>
<td>E2 (L)</td>
<td>6/7.5</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>– (R)</td>
<td>&lt;6/360</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pc</td>
<td>E3 (L)</td>
<td>6/7.5</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>E14 (R)</td>
<td>6/30</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pd</td>
<td>E4 (L)</td>
<td>6/9.5</td>
<td>Dry AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>– (R)</td>
<td>6/9.5</td>
<td>Dry AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pe</td>
<td>E6 (R)</td>
<td>6/12</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>– (L)</td>
<td>&lt;6/360</td>
<td>End-stage AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pf</td>
<td>E7 (L)</td>
<td>6/12</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>E18 (R)</td>
<td>6/120</td>
<td>End-stage AMD/Dry</td>
<td>4</td>
</tr>
<tr>
<td>Pg</td>
<td>E8 (L)</td>
<td>6/12</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>E17 (R)</td>
<td>6/120</td>
<td>End-stage AMD/Dry</td>
<td>4</td>
</tr>
<tr>
<td>Ph</td>
<td>E9 (R)</td>
<td>6/12</td>
<td>Dry AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>– (L)</td>
<td>6/15</td>
<td>Dry AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pi</td>
<td>E10 (R)</td>
<td>6/19</td>
<td>Dry AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>E15 (L)</td>
<td>6/30</td>
<td>Dry AMD</td>
<td>4</td>
</tr>
<tr>
<td>Pj</td>
<td>E11 (L)</td>
<td>6/24</td>
<td>Dry AMD</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>E16 (R)</td>
<td>6/95</td>
<td>End-stage AMD/Dry</td>
<td>4</td>
</tr>
<tr>
<td>Pk</td>
<td>E12 (L)</td>
<td>6/24</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>– (R)</td>
<td>6/240</td>
<td>End-stage AMD/Dry</td>
<td>4</td>
</tr>
<tr>
<td>Pl</td>
<td>E13 (L)</td>
<td>6/24</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>E19 (R)</td>
<td>&lt;6/360</td>
<td>End-stage AMD/Dry</td>
<td>4</td>
</tr>
<tr>
<td>Px</td>
<td>– (L)</td>
<td>6/4.8</td>
<td>Dry AMD</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>– (R)</td>
<td>6/7.5</td>
<td>Wet AMD</td>
<td>4</td>
</tr>
</tbody>
</table>

*Eyes marked with "–" were not eligible for inclusion in the study (see Supplement 5). Eyes given codes (E1, E2 etc.) met inclusion criteria, and are numbered the same as in Table 1 of main text.

*bLCVA = low contrast visual acuity; LCVA <6/240 indicates the patient could not read all letters on the largest line of the LCVA chart, "–" indicates not tested; BCVA <6/360 indicates the patient is counting fingers only. LCVA correlated very highly with BCVA (r = .95 for the 19 AMD-affected eyes meeting study inclusion criteria).
SUPPLEMENT S6. Additional procedure details for Experiment 2.

For the binocular practice phase, AMD patients were shown six practice trials presented without a time restriction, and were given feedback on whether their response was correct. These showed faces not used in the main experiment, and showed one of each expression (all veridical). The practice trials were then repeated with the restricted presentation time (5 seconds per face).

For the real experiment, patients were warned that the number of expressions would not be equal between each of the six emotions, and also to ignore the identity of the face (i.e., they would see a variety of different people, but each person may not display all of the emotions). They were also informed that similar to real-life, some of the faces are very expressive and their emotions may be easier to recognise, whereas other faces will be less expressive and may be harder to recognise and so patients should not be concerned if they could not recognise all or many of the expressions. No mention of caricaturing or image manipulation was made.

Stimuli were presented on an Apple iMac computer (screen size 68.5cm, resolution = 2560 x 1440 pixels) running OS X, using SuperLab 4.5 software. Patients were monitored for fatigue or discomfort, and offered regular breaks.

The following script contains the instructions given to AMD patients:

Show the instruction slides with BOTH eyes
- You will be looking at faces on the computer screen with one eye only and will make some decisions about them. Place the participant 40 cm from the screen. You are free to move your head around when looking at the screen, especially if you find that moving your head helps you see the faces more clearly. Please don’t move your face forward, closer to the screen. If I notice you are moving forward during the experiment, I will place you back in the correct position.
- At any time during the experiment if you need to move, stand up, stretch or have a break please let me know. Also, if you are finding the task tiring, or straining on your eyes, let me know and we can take a break. Check the participant is in a comfortable position e.g. chair height etc.
- During the experiment you will see faces on the screen that will have one of six possible emotions. I will ask you: What emotion is being expressed by this face? and you can choose from anger, disgust, fear, happy, sad and surprise (point to large-text list of emotions on card below the computer screen).
- Here are some examples of faces on the screen and I want you to tell me what emotion is being expressed by each face choosing between anger, disgust, fear, happy, sad and surprise. Present the example slides of the six different emotions and ask participants “What emotion is being expressed by this face; from anger, disgust, fear, happy, sad and surprise?”, wait for the patient to respond (there is no time restriction for slide presentation), then tell the participant what the emotion on each slide is “e.g., you are correct/incorrect, this face is expressing anger”.
- Ask the participant if they can remember what the six emotions are to check they can remember them all and present the slide with the six emotions table as a reminder.
- Now we are going to see what the experiment looks like. This time when you see the face on the screen you are going to say the emotion being expressed on the face out loud and I will enter your response into the keyboard. This time the face will only be presented on the screen for 5 seconds and then disappear. You can make your choice after the face has gone away from the screen, however it is recommended you try to choose one of the six facial
expressions as quickly and accurately as you can, you don’t have to wait the 5 seconds to decide.

- Here are some examples of faces being presented on the screen for 5 seconds. Like the last practice, you need to tell me what emotion is being expressed by each face choosing between anger, disgust, fear, happy, sad and surprise. **Present the time restricted example slides of the six different emotions and ask participants “What emotion is being expressed by this face; from anger, disgust, fear, happy, sad and surprise?”**, wait for the patient to respond, then tell the participant what the emotion on each slide is “e.g., you are correct/incorrect, this face is expressing disgust”.

- During the experiment, the number of expressions you see will not be equal between each of the six emotions, so don’t feel like you need to say each emotion an equal amount of times.

- Some of the people you see in the experiment may be expressing different emotions during the experiment and each person may not display all of the emotions, so base your response on the emotion you can see and not on the specific person.

- You might notice that the intensity of emotions across the faces varies. This is similar to real-life, for example, some people are very expressive and it is easy to recognise their emotions, whereas other people are much less expressive and therefore it is less obvious which emotion that person is expressing. That is normal, so just try to choose the emotion on each face.

- We will be measuring your accuracy during the experiment, so try to recognise the emotion on each face as best as you can.

- In the first block of the experiment you will be using your stronger eye and your weaker eye will be covered by an eye patch. This will be reversed in the second block.

- Do you have any questions? Do you feel comfortable with what the task involves? Would you like to see the introduction/practice slides again?

- We will have a break (tea/coffee) half way through the experiment, but if you need a break at any time, please let me know.

**FIRST EYE TO BE TESTED (stronger eye if both eyes tested)**

- Please cover your weaker eye now with the eye patch so you are only using your stronger eye. Your eye might take a little time to adjust. **Wait for one minute**.

- Remove the eye patch when completed and have a break.

**IF A SECOND EYE IS BEING TESTED (on a different day)...**

- Instructions and practice as for the first eye.
SUPPLEMENT S7: References for Supplementary Materials


