

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

Research School of Psychology and ARC Centre of
Excellence in Cognition and its Disorders,
The Australian National University,
Canberra, ACT, Australia

Jo Lane



Rachel A. Robbins

Research School of Psychology,
The Australian National University,
Canberra, ACT, Australia

Emilie M. F. Rohan

John Curtin School of Medical Research (JCSMR),
The Australian National University,
Canberra, ACT, Australia

Kate Crookes

Research School of Psychology and ARC Centre of
Excellence in Cognition and its Disorders,
The Australian National University,
Canberra, ACT, Australia
School of Psychological Science,
University of Western Australia, Perth, WA, Australia

Rohan W. Essex

Academic Unit of Ophthalmology, Medical School,
The Australian National University,
Canberra, ACT, Australia

Ted Maddess

John Curtin School of Medical Research (JCSMR),
The Australian National University,
Canberra, ACT, Australia

Faran Sabeti

John Curtin School of Medical Research (JCSMR),
The Australian National University,
Canberra, ACT, Australia
Discipline of Optometry and Vision Science,
The University of Canberra, Bruce, ACT, Australia
Collaborative Research in Bioactives and Biomarkers
(CRIBB) Group, Canberra, ACT, Australia

Jamie-Lee Mazlin

Research School of Psychology,
The Australian National University,
Canberra, ACT, Australia

Citation: Lane, J., Robbins, R. A., Rohan, E. M. F., Crookes, K., Essex, R. W., Maddess, T., Sabeti, F., Mazlin, J.-L., Irons, J., Gradden, T., Dawel, A., Barnes, N., He, X., Smithson, M., & McKone, E. (2019). Caricaturing can improve facial expression recognition in low-resolution images and age-related macular degeneration. *Journal of Vision*, 19(6):18, 1–22, <https://doi.org/10.1167/19.6.18>.



Jessica Irons

Research School of Psychology,
The Australian National University,
Canberra, ACT, Australia

Tamara Gradden

Research School of Psychology,
The Australian National University,
Canberra, ACT, Australia

Amy Dawel

Research School of Psychology and ARC Centre of
Excellence in Cognition and its Disorders,
The Australian National University,
Canberra, ACT, Australia

Nick Barnes

Research School of Engineering, The Australian National
University and Data61, Commonwealth Scientific and
Industrial Research Organisation,
Canberra, ACT, Australia

Xuming He

School of Information Science and Technology,
Shanghai Tech University, Shanghai, China

Michael Smithson

Research School of Psychology,
The Australian National University,
Canberra, ACT, Australia

Elinor McKone

Research School of Psychology and ARC Centre of
Excellence in Cognition and its Disorders,
The Australian National University,
Canberra, ACT, Australia

Previous studies of age-related macular degeneration (AMD) report impaired facial expression recognition even with enlarged face images. Here, we test potential benefits of *caricaturing* (exaggerating how the expression's shape differs from neutral) as an image enhancement procedure targeted at mid- to high-level cortical vision. Experiment 1 provides proof-of-concept using normal vision observers shown blurred images as a partial simulation of AMD. Caricaturing significantly improved expression recognition (happy, sad, anger, disgust, fear, surprise) by ~4%–5% across young adults and older adults (mean age 73 years); two different severities of blur; high, medium, and low intensity of the original expression; and all intermediate accuracy levels (impaired but still above chance). Experiment 2 tested AMD patients, running 19 eyes monocularly (from 12 patients, 67–94 years) covering a wide range of vision loss (acuties 6/7.5 to poorer than 6/360). With faces pre-enlarged, recognition approached ceiling and was only slightly worse than matched controls for high- and medium-intensity expressions. For low-intensity expressions, recognition of veridical expressions remained impaired and was significantly improved with caricaturing across all levels of vision loss by 5.8%.

Overall, caricaturing benefits emerged when improvement was most needed, that is, when initial recognition of uncaricatured expressions was impaired.

Age-related macular degeneration (AMD) is the most common cause of irreversible vision loss in the developed world (Bunting & Guymer, 2012; Khandhadia, Cipriani, Yates, & Lotery, 2012). It causes progressive damage to the retina, impairs central vision, and reduces visual acuity. Patients report that faces and other objects can appear blurred, distorted, and/or with missing parts (Lane, Rohan, Sabeti, Essex, Maddess, Dawel, et al., 2018; Taylor, Edwards, Binns, & Crabb, 2018).

One result of AMD is impaired facial expression recognition (Boucart et al., 2008; Johnson, Woods-Fry, & Wittich, 2017; Tejeria, Harper, Artes, & Dickinson, 2002). This can result in an inability to recognize other people's emotions and social signals in everyday life, leading patients to suffer difficulties in social interactions and contributing to social withdrawal (Lane, Rohan, Sabeti, Essex, Maddess, Dawel, et al., 2018). It is, thus, important to develop techniques that have the

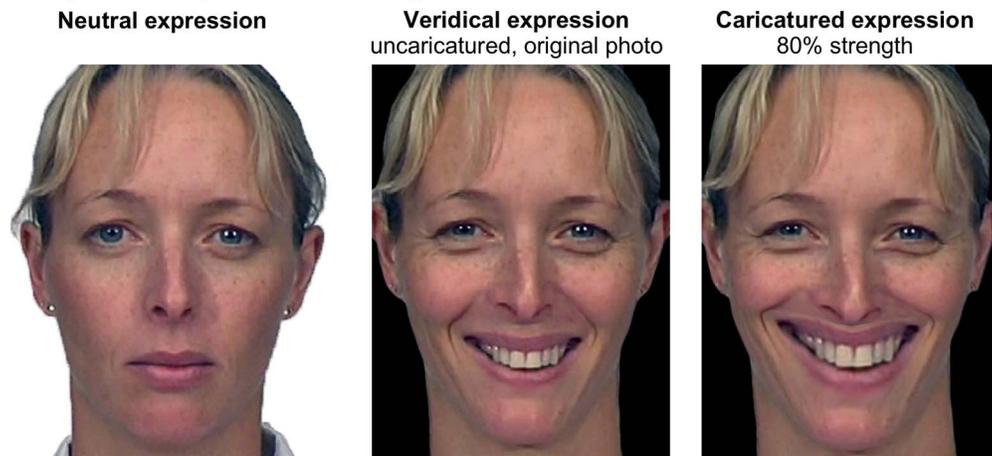
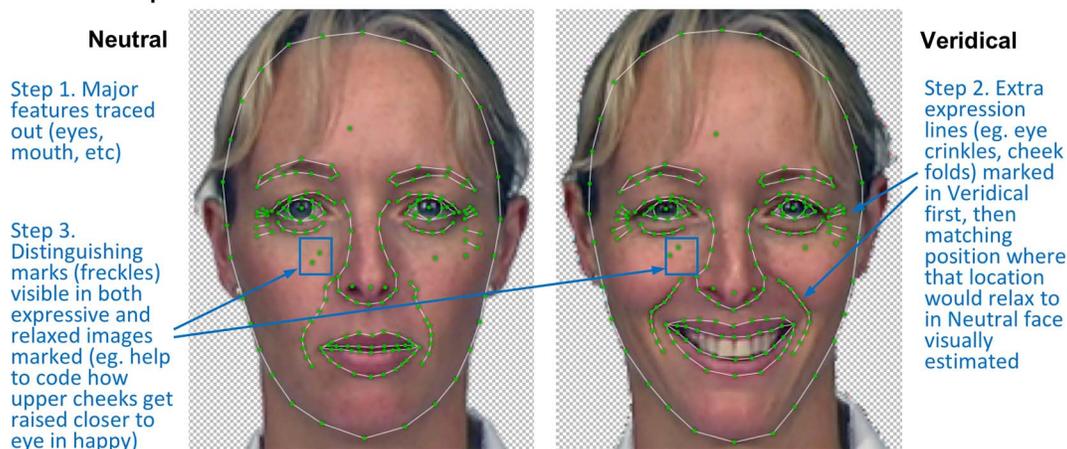
A. Example of our expression caricaturing:**B. Landmark point locations used to make the caricature:**

Figure 1. Expression caricaturing. (A) Example of our caricaturing of a happy expression. Neutral and veridical images from McLellan database (McLellan, Johnston, Dalrymple-Alford, & Porter, 2010) and published with permission from Tracey McLellan. (B) Location of the landmark points (green dots) we used to make the caricature.

potential to improve patient expression recognition ability.

One way to do this is via *image enhancement*, that is, by altering the facial image in a way that makes the face easier for the patient's brain and cognitive processes to perceive (Irons et al., 2014; van Rheede et al., 2015). To date, only one enhancement technique has been tried in AMD for face expression, namely enlargement of the image. Increasing face size improves expression recognition in AMD patients (Johnson et al., 2017; Tejeria et al., 2002). However, it does not improve it to the level of age-matched controls even for images enlarged to 21° or 44°, equivalent to seeing a real person's head 53–24 cm away (Johnson et al., 2017). Thus, there is a need to explore additional types of image enhancements.

In the present study, we focus on *caricaturing* (Figure 1) as a potential additional image enhancement. Enlargement is targeted at improving early stage visual processing (e.g., in retina to V1). Caricaturing, in contrast, is targeted at improving later stage coding of

face *shape* in mid- and high-level visual processing areas. Such shape coding occurs in regions of inferotemporal cortex sensitive to facial expression (e.g., superior temporal sulcus, fusiform gyrus; Wegrzyn et al., 2015) plus areas sensitive to general shape information in all objects (e.g., V4, lateral occipital complex; Kanwisher & Dilks, 2013; Kayaert, Biederman, Op de Beeck, & Vogels, 2005; Pasupathy & Connor, 2001). This targeting of mid- to high-level vision gives the potential for any benefits of caricaturing to be additive with early vision-targeted enhancements, such as enlargement. It also has the advantage that the caricaturing improvement is likely to be independent of the variation across individual AMD patients in the exact type, location, and severity of retinal damage and also across the individual patient variation in exact visual appearance of faces.

In a previous study, we have shown that caricaturing improves AMD patients' ability to perceive facial *identity*—that is, how one person's face differs from

another's—across a wide range of vision loss severities, including for patients with mild, moderate, and even severe vision loss (legally blind; Lane, Rohan, Sabeti, Essex, Maddess, Barnes, et al., 2018). The identity caricaturing benefits in AMD patients also paralleled those in a partial simulation of AMD that used normal vision observers shown face images with added blur (Dawel et al., 2019; Irons et al., 2014; McKone, Robbins, He, & Barnes, 2018).

Here, we ask whether caricaturing might be useful for improving poor recognition of facial *expressions* in AMD. For expression, caricaturing involves exaggerating the ways in which a particular expression (e.g., happy) differs physically from an image of the same person displaying a neutral expression (Calder, Young, Rowland, & Perrett, 1997). To make the caricature (Figure 1), multiple landmark points are assigned to the expressive version of the face (the original expression, referred to as the *veridical* image), and the matching locations are marked in the relaxed, neutral version. Morphing software is then used to exaggerate the differences between landmark locations. This exaggeration can be performed to differing degrees, resulting in different caricature strengths (Figure 2).

Caricaturing is known to improve expression perception in normal, high-resolution vision. Across both young and older adults, caricaturing can improve speed of naming the expression, sometimes improve accuracy if this is not already at ceiling for veridical expressions, and increase ratings of “how much” of the target emotion the face is displaying (Benson, Campbell, Harris, Frank, & Tovée, 1999; Calder et al., 2000; Calder et al., 1997; Leppänen, Kauppinen, Peltola, & Hietanen, 2007; Kumfor, Irish, Hodges, & Piguet, 2013; Kumfor et al., 2011).

Our present study examines whether caricaturing can improve expression perception in low-resolution vision, testing accuracy of recognizing the six so-called *basic expressions* (happy, sad, anger, disgust, fear, surprise; Ekman, 1993). Experiment 1 provides a proof of concept by testing normal vision observers—both young adults and older adults in the age range relevant for AMD—shown blurred face images (Figure 3). Experiment 1 also provides information on the range of circumstances under which caricature benefits emerge (e.g., the levels of simulated vision loss, the range of performance accuracies for veridical expressions). Experiment 2 then tests AMD patients, combining pre-enlargement of the images (Johnson et al., 2017) with caricaturing to test, first, whether caricaturing can assist expression recognition in AMD and, second, with the combination of two image enhancements, how close patient performance gets to normal vision for the patient's age group.

Finally, an additional variable included in both our experiments is the *intensity* of the original, uncarica-

tured expression. In real life, people's natural expressions vary widely in intensity. Patients' everyday social interactions would commonly include not only strong expressions (e.g., the high-intensity happy face in Figure 1), but also subtler cues to others' emotions (e.g., the low-intensity sad face in Figure 2). Previous AMD expression studies (Boucart et al., 2008; Johnson et al., 2017; Tejeria et al., 2002) have not discussed intensity nor reported the intensity of their stimuli. However, given that low-intensity expressions contain only small physical differences from neutral, we would expect that the reduced acuity in AMD is likely to result in particularly poor ability to see low-intensity expressions compared to seeing the larger physical changes present in a more intense version of an emotion.

Experiment 1: Normal vision observers shown blurred face images

The aim of our first experiment was to provide a proof of concept that caricaturing can improve expression recognition in low-resolution vision. All previous studies showing caricature benefits in expression recognition have tested the high-resolution situation, namely high-resolution face images viewed by observers with normal, high-resolution vision. Here, we test normal vision observers on low-resolution images.

The specific form of low resolution we tested was Gaussian blur. Blur is the most common feature of visual appearance reported by patients with AMD (Lane, Rohan, Sabeti, Essex, Maddess, Dawel, et al., 2018). We use the same blur simulation that we have previously shown to produce parallel results between real patients and simulations in face identity recognition (Irons et al., 2014; Lane, Rohan, Sabeti, Essex, Maddess, Barnes, et al., 2018). We tested two levels of blur (labeled “Blur 50” and “Blur 70”) as illustrated in Figure 3 as an analogy for different severities of vision loss.

The design of Experiment 1 crossed image resolution (high-resolution, Blur 50, Blur 70) with three caricature strengths: 0% (the veridical expression), 40%, and 80%. The primary research question was whether caricaturing could be found to improve accuracy of expression recognition compared to veridical for one or both levels of blur. We also examined whether caricature benefits might vary with intensity of the veridical expression (high, medium, low). Finally, we tested whether caricature benefits for blurred images remain equally strong across the full adult age range or might, for example, become weaker

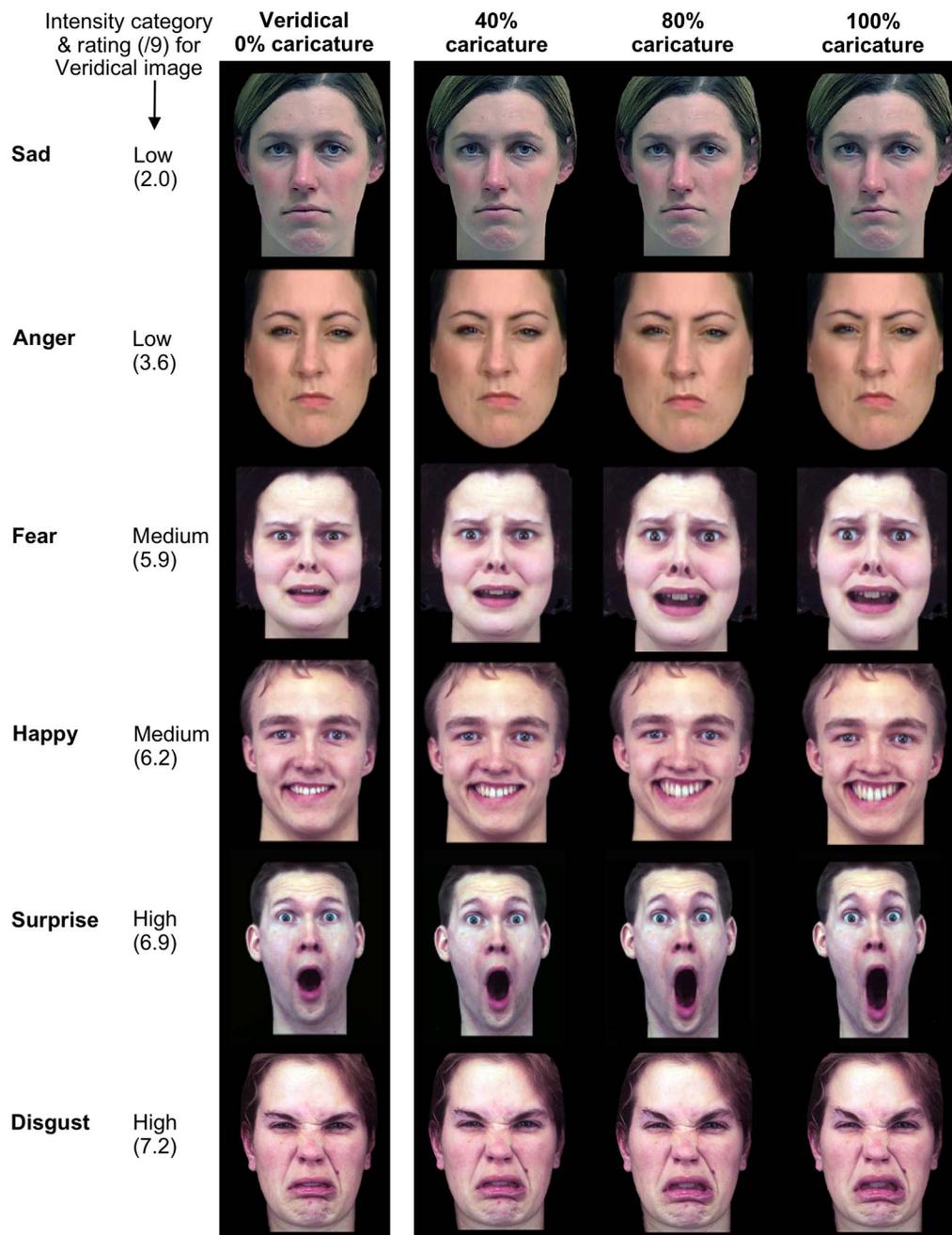


Figure 2. Example expression stimuli, selected to illustrate: the six basic expressions (Ekman, 1993) we tested: a range of expression intensities for the original (veridical) face and the caricature levels we tested (0, 40, and 80 in Experiment 1; 0, 40, 80, and 100 in Experiment 2). Numbers in parentheses give the mean intensity rating for the veridical image on scale of 1 = “weak” to 9 = “strong.” Veridical images from McLellan (sad, F009; angry, F004; McLellan et al., 2010) and KDEF databases (fear, AF16; happy, AM08; surprise, AM11; disgust, AF12; Lundqvist et al., 1998).

as the brain ages. We, thus, tested a young adult group (mean age 21 years) and an older adult group (mean age 73 years, range 65–89). The wide age range in the older adult group allowed us to evaluate, in addition to the group-mean caricature improvement, the correlation with exact age *within* this group. The rate of AMD increases with age, and thus, it is of interest

to know whether, for example, caricature benefits might be present at the younger end of the AMD-relevant age range (mid-60s; Kumfor et al., 2013; Kumfor et al., 2011) but perhaps fall off in very elderly participants (e.g., over 80 years of age).

The task was six-alternative, forced-choice recognition (anger, happy, sad, fear, disgust, surprise). Note

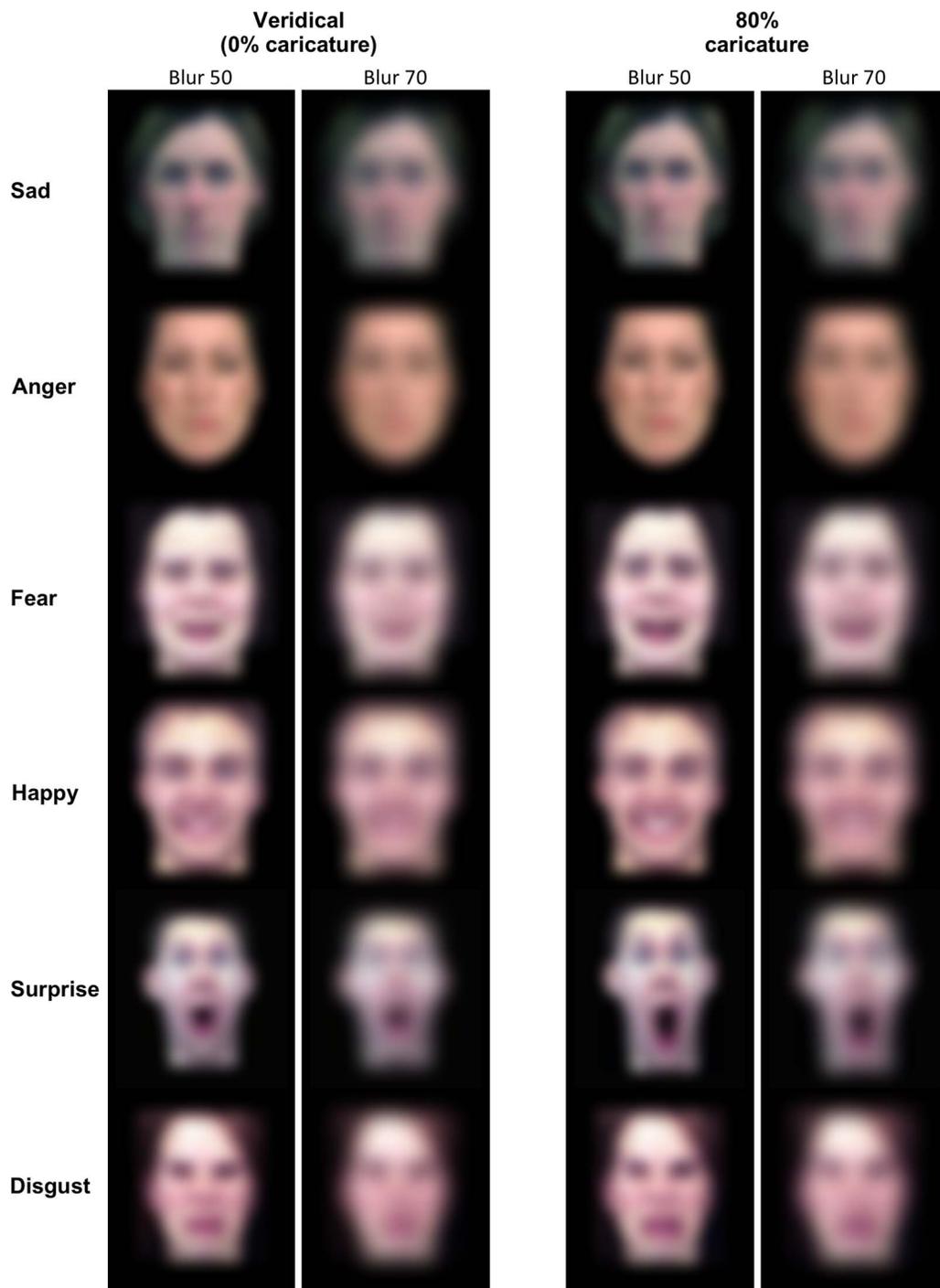


Figure 3. Blur levels added in Experiment 1.

we do not necessarily expect caricaturing to always improve recognition in this task, particularly when accuracy approaches floor (chance) or ceiling (Calder et al., 1997). Concerning what value represents “ceiling,” our stimuli are face only and are shown without the contextual information that assists emotion recognition in everyday life (e.g., body language, scene context, knowledge of past history of events; Aviezer, Trope, & Todorov, 2012). For face-only

stimuli, mean recognition accuracy averaged over the six basic expressions rarely exceeds 85% even for “gold standard” expression sets viewed with perfect vision (Palermo & Coltheart, 2004). Our primary interest, then, is in whether caricature benefits emerge when veridical recognition is impaired by blur (or other factors) to noticeably below the best possible of 85% correct while remaining sufficiently above the chance value of 17%.

Intensity category	Intensity rating for veridical ^a , <i>M</i> (<i>SD</i>) [range]	Number of face items						
		Total	Anger	Happy	Sad	Surprise	Fear	Disgust
Low	3.34 (0.84) [1.92–4.76]	27	8	2	10	1	2	4
Medium	5.74 (0.48) [4.80–6.36]	27	2	6	4	7	5	3
High	7.09 (0.58) [6.40–8.36]	28	4	6	6	2	4	6

Table 1. Properties of low-, medium-, and high-intensity face subsets. ^aIntensity rating task was “How intense does this emotional expression look to you?” with response scale running from 1 = “weak” to 9 = “strong.”

Method

Participants

Young adults were tested in groups as part of undergraduate laboratory classes in psychology at the Australian National University. Participants included in the analysis ($n = 45$; age $M = 20.6$, range = 19–24; 33 female, 12 male) confirmed they had normal or corrected-to-normal vision, were wearing their usual glasses or contacts for screen viewing when relevant, wished to have their data kept for research purposes, considered their data to be “valid” (i.e., the computer didn’t crash during the experiment, they weren’t abnormally tired or sick), did not have autism spectrum disorder (ASD, which can impair facial expression recognition), and were Caucasian (the same race as the face stimuli to avoid other-race effects on expression recognition; Elfenbein & Ambady, 2003).

Older adults ($n = 29$; age $M = 73.3$, range = 65–89; 23 female, 6 male) were tested individually. Their vision was tested in each eye using an ETDRS chart viewed at the experimental testing distance (60 cm) and wearing their normal glasses for that viewing distance. All had monocular best-eye acuity of at least 6/9 ($n = 10$ were 6/6, $n = 8$ were 6/7.5, and $n = 11$ were 6/9), and all reported their binocular eyesight was fine for recognizing people and seeing movies and plays in everyday life. All were Caucasian and none reported having ASD. As a quick screening tool to exclude dementia, all passed the Mini Mental State Exam (score >25 ; Folstein, Folstein, & McHugh, 1975) except one participant who ran out of time to do this test because she was going to her Latin class.

Participants gave informed written consent prior to participating. The experiment was approved by the Australian National University Human Research Ethics Committee and complied with the Declaration of Helsinki.

Stimuli

Veridical expression faces and corresponding neutral faces needed to make caricatures: Veridical expressions and corresponding neutral images were taken from four databases: Karolinska Directed Emotional Faces (KDEF; Lundqvist, Flykt, & Öhman, 1998), NimStim (Tottenham et al., 2009), McLellan (McLellan, John-

ston, Dalrymple-Alford, & Porter, 2010), and Gur (Gur et al., 2002). The *veridical* expressions were 82 color, front-view photographs, showing anger (14 images), disgust (13), fear (11), happiness (14), sadness (20), or surprise (10). Images came from a total of 48 Caucasian young adults (24 females, 24 males). Selection of items (an uneven number across emotions) was based on meeting multiple inclusion criteria: good quality photographs, availability of neutral expression reference image showing the same person, availability of matched mouth position across veridical and neutral (e.g., for mouth open anger, we required a mouth open neutral because using mouth closed introduces morphing artifacts into the caricatures), good labeling accuracy (as provided in the original database articles, i.e., for example, most people shown a face labeled “anger” in the database agreed it did, indeed, display anger), and covering a range of expression intensities. Faces were placed on a standard-sized black background and images cropped to show the region from chin to approximately the hairline (see examples in Figure 2), using Adobe Photoshop Elements 12 software.

Intensity of veridical expressions: Table 1 details the division of veridical faces into three intensity categories. This was based on data from an intensity rating experiment described in Supplementary File S1, Supplement S1 ($n = 25$ young adults), which was used to rank order the 82 veridical faces and divide them into thirds into *low-*, *medium-*, and *high-*intensity sets.

Expression caricaturing: Caricatures were created using Abrosoft Fantamorph 5.3.0. Multiple landmark points were manually placed on each veridical image (Figure 1B), tracing out the shape of all major features (eyes, nose, mouth, eyebrows, hairline, and face outline, including cheek and chin shape) plus any extra expression-related lines. For particular images, extra lines could include wrinkle lines across the top of the nose if these were visible in a disgust face or upward-curving lines in the forehead between the eyes in sad. Matching locations were then marked on the corresponding neutral expression image. For major features, this is straightforward (i.e., a marker dot at the left corner of a smiling mouth is paired with a marker dot at the left corner of a neutral mouth). For the extra expression-related lines, the lines often disappear in

neutral; we marked the paired location as being our best visual estimate of where the expressive face location would relax to in the neutral expression. Where the individual person had additional distinguishing features (e.g., moles visible in both the veridical and the neutral version), some were also marked to help match locations of the same piece of skin across the expressive and neutral versions.

For expressions displaying teeth, these were often not visible in open-mouthed neutral versions. We, thus, matched landmark locations based on the inside line of the lips with no landmarks around the teeth. This results in exaggeration of the size of the teeth in the caricatures while maintaining the proportion of tooth size to size of gap between the lips (see happy example in Figure 1A). We judged this to be the best way to caricature the apparent strength of the emotion displayed; also note that the alternative of not caricaturing the teeth at all (i.e., keeping them the same size as in the veridical version) often led to a very peculiar appearance (e.g., an impression of tiny teeth in a huge mouth for expressions with a gap between top and bottom teeth).

The final number of landmark points was approximately 140–230 points per face (varying with different expressions and different individual models). Caricatures were then extracted from Fantamorph at 0% (veridical), 40%, and 80% strengths; 100% would indicate a doubling of the differences between veridical and neutral landmark point locations. Only shape information was caricatured (in morphing software language, caricaturing was applied only to *warp* and not *fade* functions); this is because, in the real world, patients would see faces varying in lighting, and caricaturing nonshape information exaggerates lighting information that, in some cases, could be misleading as to the expression displayed (see General discussion for further consideration of this issue).

Addition of blur: We took the 246 high-resolution images described above (82 veridical images, 82 images at 40% caricature, 82 images at 80% caricature) and rendered each in two levels of blur illustrated in Figure 3, labeled “Blur 50” (less extreme) and “Blur 70” (more extreme). Testing two levels of blur was designed to capture the idea that, in AMD, the severity of vision loss varies across patients.

The specific blur formula used was Marmor and Marmor’s (2010) formula for blurring perceived in peripheral vision; this is of relevance to disorders producing central vision loss, such as AMD (although note that some patients might rely on islands of intact retina in central vision rather than peripheral vision and also that blur does not provide a complete simulation of all AMD patients’ experience; Lane, Rohan, Sabeti, Essex, Maddess, Dawel, et al., 2018). We applied uniform spatial blur across the image by

reducing the contrast of spatial frequencies higher than a given threshold (with threshold set lower for higher blur levels, using a Gaussian kernel filter of size defined by the cutoff frequency; Supplement S2 provides details). The labels of our two conditions (“Blur 50” and “Blur 70”) are somewhat arbitrary, but use numbering consistent with that we have previously employed in face identity studies (in which we tested up to “Blur30”; Dawel et al., 2019; Irons et al., 2014; McKone et al., 2018): Specifically, the amount of blur is as the Marmor and Marmor formula gives for peripheral viewing of a face assumed to subtend 18.11° along the horizontal (equivalent to a real person viewed at 54 cm) at 50° eccentricity (“Blur 50” condition) or 70° eccentricity (“Blur 70” condition). Note that the same amount of blur results in the Marmor and Marmor model for a smaller face seen less far into the periphery.

Procedure

On each trial, the face appeared at screen center until response. Participants responded to the question “What emotion do you think is being expressed by this face?” by using the mouse to select one of six options via on-screen buttons (anger, disgust, fear, happy, sad, surprise). Two of the older adult participants responded verbally and had their responses entered by the experimenter (e.g., due to not being confident using computers). Viewing was binocular. Interval between trials was 300 ms. Target viewing distance was 60 cm (no chin rest was used), making face images approximately 8.6° – 10.9° tall (covering the region from just below the chin to just above the hairline), which is equivalent to viewing a real-world person from 128 cm away (calculated using the fact that average real head size is 22 cm; Farkas, Hreczko, & Katic, 1994; McKone, 2009); note that size varies across images due to natural variation in aspect ratio of faces and also, in some cases, due to caricaturing (e.g., caricaturing surprise makes the face longer vertically; see example in Figure 2). Importantly, this size is large enough to achieve maximum performance for normal vision observers (i.e., the limitation on further accuracy improvement is not image size or resolution, but rather the use of context-free face-only stimuli).

Each participant was shown 738 trials (82 images \times 3 caricature levels \times 3 blur levels). Resolution, caricature, and intensity conditions were randomly intermixed across the full set of trials with the 738 conditions/face items presented in a different random order for each participant. This randomization ensured that caricature (and blur) effects were not confounded with any practice benefits due to reusing the same face items across conditions (i.e., the average position in the list

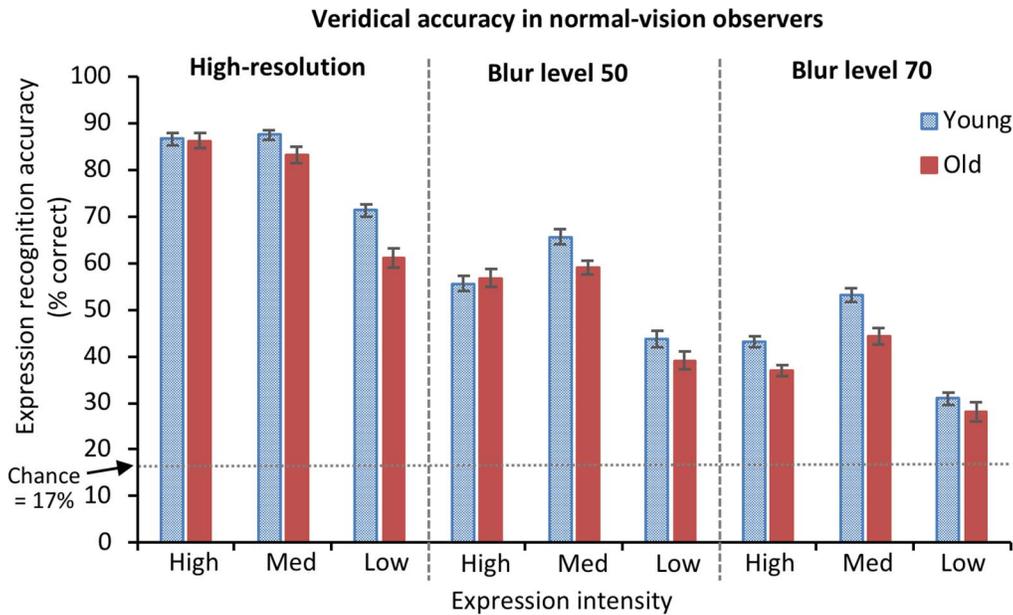


Figure 4. Experiment 1 results: mean expression recognition accuracy for veridical (uncaricatured) faces in normal vision observers. Error bars show ± 1 SEM.

was equal across the three caricature levels for each of the 82 face items).

Trials were divided into six blocks of 123 trials with a short rest break after each block. Supplement S3 details computer equipment and software.

Results

Recognition of veridical (uncaricatured) expressions

We first describe recognition accuracy for the original (veridical) expression stimuli in order to confirm some basic expectations that recognition performance should worsen with increasing blur (Irons et al., 2014), recognition should be more difficult for low-intensity than high-intensity expressions (Palermo & Coltheart, 2004), and there should be only a small reduction in overall expression recognition ability with aging (c.f. $\sim 5\%$ reduction in high-resolution image studies included in meta-analysis of Ruffman, Henry, Livingstone, & Phillips, 2008). Results in Figure 4 show that all these predictions were supported.

A 3 (blur level) \times 3 (intensity level) \times 2 (age group) ANOVA revealed a significant main effect of blur, $F(1.664, 119.841)^1 = 980.511$, $MSE = 106.301$, $p < 0.001$, in which, as image resolution was decreased, expression accuracy dropped substantially (e.g., $\sim 86\%$ correct for high-resolution down to $\sim 40\%$ for Blur70 for high-intensity expressions; Figure 4). There was also a main effect of intensity level, $F(2, 144) = 377.452$, $MSE = 59.793$, $p < 0.001$ with low-intensity expressions being the most poorly recognized and medium-intensity expressions similar to high intensity (Figure 4). Finally,

there was a significant but small main effect of age, $F(1, 72) = 11.153$, $MSE = 319.293$, $p = 0.001$ with the older adult group having, on average, 4.7% lower accuracy. (Another finding not of direct relevance to our research questions was an interaction showing stronger age-related decline for lower intensity than higher intensity expressions; see Supplement S4 for details.)

Caricature benefits

Having established that basic results were as expected, our key research question was whether caricaturing improved expression recognition relative to veridical faces and, if so, by how much and for what range of veridical “baseline” values. Results are shown in Figure 5, which plots performance across the three caricature strengths in each blur, intensity, and age group condition.

A four-way ANOVA revealed a significant main effect of caricature, $F(1.754, 126.275) = 55.064$, $MSE = 37.934$, $p < 0.001$, with higher caricature strengths producing better expression recognition accuracy. The size of the caricature benefit interacted significantly with blur, $F(4, 288) = 6.211$, $MSE = 35.066$, $p < 0.001$, and intensity, $F(4, 288) = 3.768$, $MSE = 32.679$, $p = 0.005$. Figure 5 suggests that, rather than these interactions reflecting anything theoretically interesting, variation in caricature benefit was instead related simply to the level of initial recognition accuracy for the veridical faces. Specifically, caricaturing did not produce any benefits in conditions in which veridical recognition was already very good, i.e., the four conditions in which veridical accuracy was 83% or more. However, benefits emerged when veridical

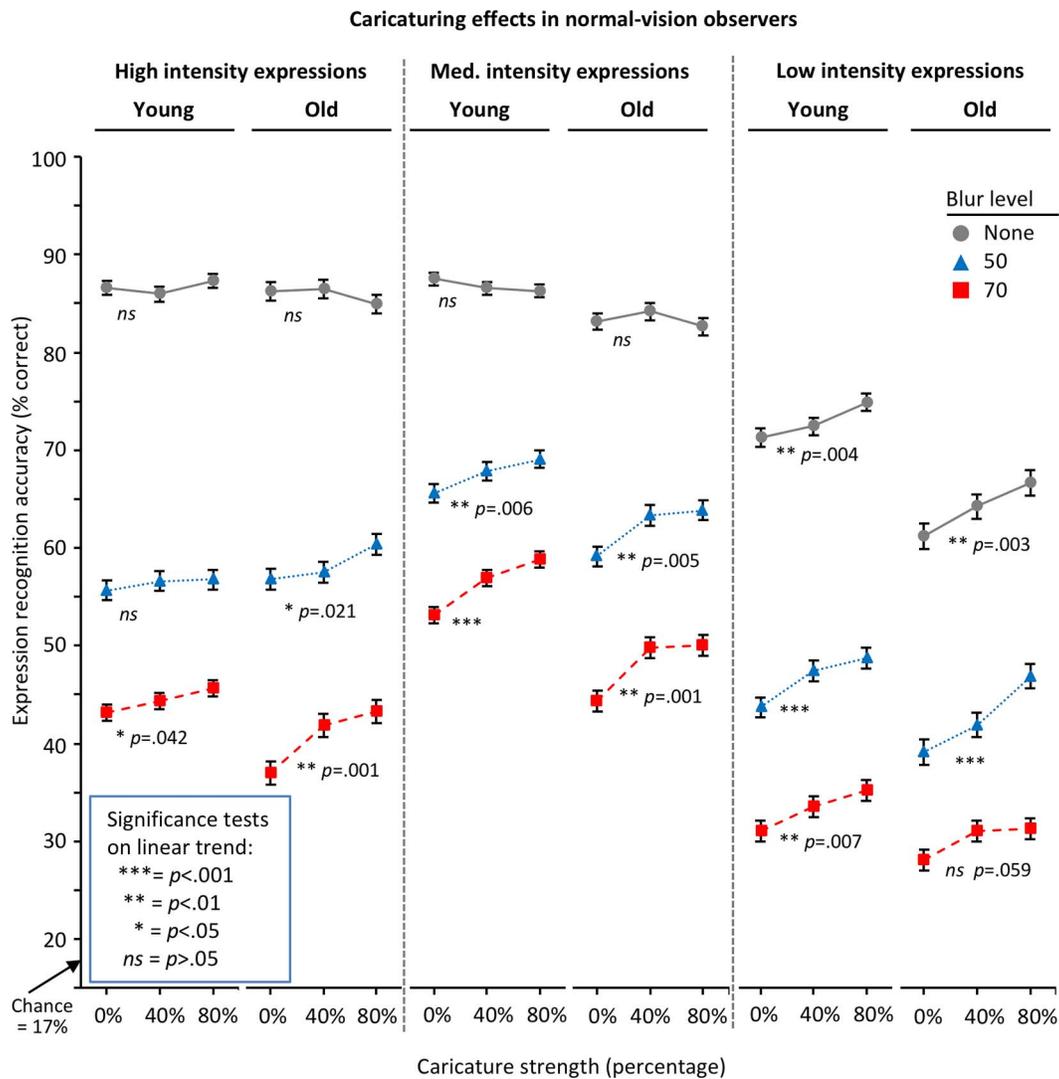


Figure 5. Caricature effects for normal vision observers in Experiment 1 for young and older adults. p = significance value for linear trend across the three caricature levels. Error bars are the equivalent of ± 1 SEM for the repeated-measures comparison of caricature strengths (calculated separately within each blur, age group, and intensity condition). Intensity refers to the intensity of the veridical (uncaricatured) expression.

performance was poorer. The remaining 14 conditions in Figure 5 had veridical accuracy of 71% or less, and of these, 12 showed a significant caricature improvement (10 at $p < 0.007$ or better; p values for linear trend analysis are summarized in Figure 5 with full statistics in Table 2), and all showed an effect in the correct direction. Overall, Figure 5 shows that significant caricature-related improvements in expression recognition could be found for all intensities and all blur levels and for both age groups as long as veridical performance fell in a range between 71% and 31% correct (in the context that chance is 17%).

Of particular relevance to potential translation to AMD patients is the question of whether caricature benefits might reduce with aging. This was not the case. The four-way ANOVA showed no interactions in-

volving caricaturing and age (no two-way caricature \times age, $p > 0.1$; no three-way caricature \times age \times blur, $p > 0.3$; no three-way caricature \times age \times intensity, $p > 0.7$; no four-way interaction, $p > 0.6$). Thus, as can be seen in Figure 5, the caricature benefits were as large in the older adult group (M age = 73 years) as in the young adult group. In addition, within the older adult group, we examined correlations with exact age (in years) across our full older adult range of 65 to 89 years. We calculated the caricature benefit for each individual participant as the accuracy for 80% caricature strength minus the accuracy for veridical. Exact age did not correlate with caricature benefit with no significant correlation for any intensity or resolution condition and, most importantly, no consistent direction of trend (Table 3).

Intensity	Caricature strength condition			Linear trend			Caricature improvement ^b
	0% veridical	40%	80%	F ^a	MSE	<i>p</i>	80% veridical
A. Young adults (<i>n</i> = 45)							
No blur							
All	81.8 (1.0)	81.7 (0.9)	82.9 (0.7)	1.994	11.964	0.165	1.0 (0.73)
Low	71.3 (1.3)	72.4 (1.4)	74.9 (1.3)	9.326	31.640	0.004	3.6 (1.2)
Medium	87.5 (1.1)	86.6 (1.0)	86.3 (0.9)	1.642	20.888	0.207	−1.2 (1.0)
High	86.6 (1.4)	86.0 (1.1)	87.3 (1.1)	0.361	31.772	0.551	0.7 (1.2)
Blur 50							
All	55.0 (1.3)	57.3 (1.5)	58.2 (1.2)	25.122	9.004	< 0.001	3.2 (0.6)
Low	43.7 (1.7)	47.4 (1.6)	48.7 (1.4)	16.666	34.030	< 0.001	5.0 (1.2)
Medium	65.6 (1.6)	67.8 (1.8)	69.1 (1.3)	8.422	31.924	0.006	3.5 (1.2)
High	55.6 (1.6)	56.6 (1.8)	56.7 (1.5)	0.849	32.706	0.362	1.1 (1.2)
Blur 70							
All	42.4 (1.1)	44.9 (1.2)	46.5 (1.1)	24.381	15.659	< 0.001	4.1 (0.8)
Low	31.0 (1.3)	33.6 (1.4)	35.1 (1.5)	8.029	47.457	0.007	4.1 (1.5)
Medium	53.1 (1.5)	57.0 (1.7)	58.8 (1.4)	23.589	31.611	< 0.001	5.8 (1.2)
High	43.1 (1.3)	44.3 (1.3)	45.6 (1.3)	4.368	33.228	0.042	3.6 (1.2)
B. Older adults (<i>n</i> = 29)							
No blur							
All	77.0 (1.4)	78.4 (1.3)	78.2 (1.3)	3.830	5.630	0.06	1.2 (0.6)
Low	61.2 (2.0)	64.2 (1.8)	66.7 (1.7)	10.936	39.987	0.003	5.5 (1.7)
Medium	83.1 (1.7)	84.2 (1.7)	82.6 (1.7)	0.185	20.441	0.67	−0.5 (1.2)
High	86.2 (1.7)	86.5 (1.6)	85.0 (2.0)	1.142	19.258	0.294	−1.2 (1.15)
Blur 50							
All	51.7 (1.4)	54.3 (1.6)	57.1 (1.4)	29.337	14.097	< 0.001	5.3 (1.0)
Low	39.1 (1.8)	41.9 (2.2)	46.9 (1.6)	22.358	39.362	< 0.001	7.8 (1.6)
Medium	59.1 (1.5)	63.3 (1.9)	63.9 (1.8)	9.455	34.243	0.005	4.7 (1.5)
High	56.8 (1.9)	57.5 (2.2)	60.3 (1.8)	5.971	30.977	0.021	3.6 (1.5)
Blur 70							
All	36.5 (1.2)	40.9 (1.0)	41.5 (1.3)	21.610	17.372	< 0.001	5.1 (1.1)
Low	28.1 (2.1)	31.0 (1.5)	31.3 (1.8)	3.882	38.078	0.059	3.2 (1.6)
Medium	44.3 (1.7)	49.8 (1.6)	50.1 (1.8)	13.850	34.581	0.001	5.7 (1.5)
High	36.9 (1.1)	41.9 (1.4)	43.2 (1.6)	14.325	39.930	0.001	2.5 (1.2)

Table 2. Experiment 1: Caricature effects on expression recognition accuracy (percentage correct choice as anger, fear, happy, surprise, sad, or disgust) in normal vision young and older adult groups as a function of blur level and veridical expression intensity expressed as *M* (*SE*) and including statistic for the linear trend. ^aDegrees of freedom on linear trend are young adults (1, 44), older adults (1, 28). ^bThe *p* value for caricature improvement is the same as that reported for the linear trend (when there are three equally spaced levels of caricature strength; *p* for the linear trend across the three equals *p* value for a *t* test comparing the extreme ends). Bolded results highlight *p* values <0.005 and corresponding significant caricature improvement.

Finally, we consider the size of the caricature benefit. Overall, excluding the four conditions in which veridical performance exceeded 80% correct, we found the size of the caricature improvement (80% caricature strength *minus* veridical, Table 2) averaged 4.3% including both age groups (range across conditions = 1.1%–7.8%) and 4.7% specifically in older adults.

Discussion

Results of Experiment 1 provide proof of concept that caricaturing can be used to improve expression recognition when faces are seen in low resolution, in

this case, blurred images viewed by normal vision observers. Additionally, there was no reduction in caricature benefit with aging, implying that brains even of people in their 80s retain the shape and expression coding mechanisms necessary to produce caricature benefits. Once veridical recognition fell below approximately 72%, the size of the caricature benefit averaged a 4.7% increase in accuracy in older adults and remained broadly stable across changes in intensity of the original expression and across changes in severity of blur (simulating, in patients, degree of vision loss). Notably, caricaturing benefits emerged when improvement was most needed, that is, when initial recognition of uncaricatured expressions was impaired. These

Condition	<i>r</i>	<i>p</i>
All intensities, no blur	−0.232	0.227
All intensities, blur 50	0.121	0.531
All intensities, blur 70	0.027	0.890
Low intensity, no blur	0.105	0.589
Low intensity, blur 50	0.149	0.441
Low intensity, blur 70	−0.054	0.782
Medium intensity, no blur	−0.226	0.239
Medium intensity, blur 50	0.118	0.541
Medium intensity, blur 70	−0.017	0.930
High intensity, no blur	−0.288	0.130
High intensity, blur 50	−0.042	0.827
High intensity, blur 70	0.118	0.543

Table 3. Experiment 1: Correlations between older adults' exact age and their caricature benefit (i.e., 80% strength *minus* veridical).

results are encouraging for potential usefulness of expression caricaturing in AMD.

Experiment 2: AMD patients

In Experiment 2, we tested patients with AMD. All face images were now high resolution with patients seeing these in various degrees of low resolution as determined by their level of vision impairment (i.e., due to their degree of retinal damage).

Using the same face stimuli as in Experiment 1, patients were tested on the three veridical intensity levels (high, medium, low) now crossed with four caricature strengths (0%, 40%, 80%, and 100%). The 100% caricature condition was included in case stronger exaggeration might further improve patients' expression recognition as compared to the maximum 80% strength used in Experiment 1. Note, however, there is no guarantee of further improvement because, at some point, caricatured expressions begin to look noticeably weird (Mäkäräinen, Kätsyri, & Takala, 2014), which could lead to a plateau or a turnaround to produce *worse* recognition accuracy beyond some ideal caricature strength. We also increased the size of the stimuli relative to Experiment 1. The size used in Experiment 1 is large enough to achieve maximum performance for face-only stimuli in normal vision observers, but with impaired vision in AMD, increasing the stimulus size improves performance (Johnson et al., 2017; Tejeria et al., 2002). Given that size increase is the easiest image enhancement technique to implement practically, any future real-world applicability to patients would always include enlargement as the first step.

Our primary question was whether, for the pre-enlarged faces, caricaturing could improve expression

recognition in AMD—that is, whether recognition accuracy in the 80% (or perhaps 100%) caricature strength condition was better than veridical—for one or more levels of initial expression intensity. Additionally, an important question was whether any caricature benefits are found across a broad range of residual visual acuities or instead might be limited only to patients with, say, relatively mild levels of vision loss for whom enlarged faces can still be seen with some degree of clarity. Arguing against this possibility, Experiment 1 found that even severely blurred expressions (Figure 3) benefitted from caricaturing; also, for *identity* caricaturing, in Lane, Rohan, Sabeti, Essex, Maddess, Barnes, et al. (2018), we reported benefits in several AMD patients with moderate vision loss (defined as acuities poorer than 6/19 using the World Health Organization, 2015, criteria) and severe vision loss (legally blind, defined as poorer than 6/60). This suggests that there is also potential for expression caricaturing to be valuable for patients even with moderate-to-severe vision loss. Across the full range of vision loss, we tested 19 AMD-affected eyes (from 12 patients) and analyzed eyes categorized into a group with mild vision loss ($n = 9$ eyes) and a group with moderate-to-severe vision loss ($n = 10$ eyes). Finally, we examine how the size of the caricature improvements in AMD compares to that in a set of age- and sex-matched normal vision controls extracted from the older adult participants in Experiment 1.

Method

Participants

AMD patients and eyes: Experiment 2 participants were 12 AMD patients (8 females; age $M = 81.4$ years, range 67–94), diagnosed by a qualified ophthalmologist as having AMD in at least one eye. To be eligible, patients had to be Caucasian to match the race of the face stimuli. Any with a diagnosis of dementia were excluded, and patients also had to demonstrate good ability to comprehend task instructions and display no evidence of dementia over the several hours during which they interacted with the experimenter.

Recruitment targeted eyes covering the full range of vision loss severity (Table 4). Best corrected visual acuity (BCVA) ranged from 6/7.5 to poorer than 6/360. We analyze the 19 individual eyes, tested monocularly, that met inclusion criteria. The first inclusion criterion was that the eye had to have AMD and no other diagnoses; note that clinically nonsignificant visual opacity was allowed. Additionally, there were separate inclusion criteria applied at the top and bottom end of vision ability. Given that image-enhancement technology is of interest only when ability is poorer than normal vision, at the top end, we included only eyes with relevant functional vision loss. This was defined as

Eye code ^a (left or right eye)	Visual acuity (BCVA) ^b	Diagnosis AMD type	Patient code ^{c,d} (sex, age)
Mild vision loss			
E1 (R)	6/7.5	Early AMD	Pa (M, 86)
E2 (L)	6/7.5	Wet AMD	Pb (M, 72)
E3 (L)	6/7.5	Wet AMD	Pc (M, 81)
E4 (L)	6/9.5	Dry AMD	Pd (F, 79)
E5 (L)	6/9.5	Wet AMD	Pa
E6 (R)	6/12	Wet AMD	Pe (F, 70)
E7 (L)	6/12	Wet AMD	Pf (F, 78)
E8 (L)	6/12	Wet AMD	Pg (F, 93)
E9 (R)	6/12	Dry AMD	Ph (F, 86)
Moderate vision loss			
E10 (R)	6/19	Dry AMD	Pi (M, 79)
E11 (L)	6/24	Dry AMD	Pj (F, 92)
E12 (L)	6/24	Wet AMD	Pk (F, 94)
E13 (L)	6/24	Wet AMD	Pl (F, 67)
E14 (R)	6/30	Wet AMD	Pc
E15 (L)	6/30	Dry AMD	Pi
Severe vision loss			
E16 (R)	6/95	End-stage AMD	Pj
E17 (R)	6/120	End-stage AMD	Pg
E18 (R)	6/240	Wet AMD	Pf
E19 (R)	<6/360	End-stage AMD	Pl

Table 4. The 19 AMD-affected eyes meeting inclusion criteria, ordered by severity of vision loss (best corrected visual acuity) and corresponding patient information. ^aL = left eye (i.e., OS, ocular sinister), R = right eye (i.e., OD, oculus dextrus). ^bBCVA measured by qualified orthoptist using a retro-illuminated logMAR chart conforming to the Early Treatment Diabetic Retinopathy Study (ETDRS) standard format. Cutoff values for vision loss categories from ICD-10 criteria³⁰: *mild* = BCVA 6/6 to 6/18; *moderate* = poorer than 6/18 to 6/60; *severe* = poorer than 6/60. For context, in Australia, worse than 6/12 binocular vision results in loss of standard drivers' license, and 6/60 or worse is legally blind. ^cSeven patients had two eyes meeting inclusion criteria; five (Pb, Pd, Pe, Pk, Ph) had only one. Supplement S1 gives more complete vision data, including for eyes not meeting inclusion criteria. M = male, F = female. ^dA few of the present patients also participated in our previous identity-caricaturing study (Lane, Rohan, Sabeti, Essex, Maddess, Barnes, et al., 2018). Patient codes and eye numbers used here do not correspond with that article (i.e., eyes are ordered by vision loss severity within each article independently).

having visual acuity (BCVA) worse than 6/6 and expression recognition performance for veridical faces below normal vision ceiling levels (as determined for our stimuli from the mean for young adults on high-resolution images in Experiment 1). At the bottom end, we did not test any eyes in which vision was so poor that the patient reported he or she could not see the face stimuli (e.g., just a vague blob where the computer screen was). Supplement S5 provides additional details. The reason for monocular testing was to provide, when possible, data on both eyes independently (noting the

two eyes in AMD often have different visual acuity) and, thus, maximize efficiency of testing by minimizing the number of patients we needed to recruit.

Recruitment was via the Canberra Hospital Department of Ophthalmology and private ophthalmologist's rooms using a study brochure and/or personal approach while patients were waiting for their consultation; a radio interview promoting the study; and a letter sent to all local area AMD patients on the Macular Disease Foundation Australia mailing list.

Duration of participation was 2–6 hr for the expression recognition experiment (time to test a single eye ranged from 1 to 4 hr) plus 1.5 hr for vision assessment (Supplement S5). Individual sessions were <2 hr, to minimize fatigue. Patients were reimbursed for travel. Participants gave informed written consent after explanation of the nature and possible consequences of the study. Research methods adhered to the Declaration of Helsinki and were approved by the Australian National University and ACT Health Human Research Ethics Committees.

Age- and sex-matched controls (subset from Experiment 1): We also include results for a set of age- and sex-matched controls for the AMD patients. These were a subset of 12 of the older adult group extracted from Experiment 1, selected to match the patients as closely as possible. Our AMD patients were older, on average, than the full Experiment 1 group, and so we extracted controls mostly from the upper half of the Experiment 1 age range. Mean age for the controls was 79 years (range 70–89) as compared to 81 years for the AMD patients. There were eight females, the same as for the AMD patients.

Stimuli

Stimuli were identical to Experiment 1 except that, for AMD patients in Experiment 2, (a) all stimuli were high-resolution images, and (b) we included a 100% caricature strength version of each face expression extracted from Fantamorph. We did not test caricatures stronger than 100% because the images showed morphing artifacts.

Procedure

On each trial, the face appeared at screen center for five seconds. Patients were asked "What emotion is being expressed by this face?" with options read aloud and also shown in large print on a card under the screen (anger, disgust, fear, happy, sad, surprise). Patients responded verbally. The experimenter entered the response. Interval between trials was 300 ms.

Target viewing distance was 40 cm, making face images approximately 17.1° vertical, equivalent to viewing a real-world person from 58 cm away. Patients

Participant/eye group	Caricature strength condition				Linear trend (excluding 100%)			Caricature improvement ^b
	0% veridical	40%	80%	100%	F ^a	MSE	p	80% veridical
A. AMD patients (<i>n</i> = 19 eyes)								
All intensities	67.2 (2.2)	68.6 (2.7)	69.6 (2.7)	67.1 (2.4)	9.366	5.876	0.007	2.4 (0.79)
Low intensity	48.3 (3.3)	53.2 (3.4)	54.2 (3.5)	52.7 (3.7)	15.607	20.817	0.001	5.8 (1.5)
Medium intensity	74.9 (1.8)	75.2 (2.6)	75.7 (2.0)	72.5 (2.2)	0.229	31.887	0.638	0.9 (1.8)
High intensity	78.1 (2.8)	77.0 (2.7)	78.7 (3.2)	75.7 (2.5)	0.128	23.571	0.725	0.6 (1.6)
B. Mild vision loss (<i>n</i> = 9 eyes)								
All intensities	69.6 (2.7)	71.8 (3.3)	71.1 (3.9)	68.6 (3.3)	1.137	8.789	0.317	1.5 (1.4)
Low intensity	53.1 (3.9)	55.3 (4.1)	58.2 (4.8)	56.6 (5.2)	6.345	18.766	0.036	5.1 (2.0)
Medium intensity	78.0 (1.6)	80.7 (3.4)	75.5 (3.3)	73.5 (3.2)	0.970	28.292	0.354	−2.5 (2.5)
High intensity	77.4 (3.5)	79.2 (3.3)	79.2 (4.8)	75.4 (3.4)	0.493	29.098	0.502	1.8 (2.5)
C. Moderate + severe vision loss (<i>n</i> = 10 eyes)								
All intensities	65.1 (3.5)	65.7 (4.1)	68.4 (3.8)	65.7 (3.5)	16.621	3.142	0.003	3.2 (0.79)
Low intensity	44.1 (5.0)	51.3 (5.5)	50.6 (5.1)	49.3 (5.2)	8.580	24.482	0.017	6.5 (2.2)
Medium intensity	72.0 (2.9)	70.4 (3.2)	75.9 (2.5)	71.7 (3.1)	2.702	27.987	0.135	3.9 (2.4)
High intensity	78.8 (4.4)	75.0 (4.3)	78.2 (4.6)	75.9 (3.9)	0.072	19.859	0.794	−0.05 (2.0)
D. First-tested eyes (<i>n</i> = 12 eyes)^c								
All intensities	68.1 (3.1)	70.3 (3.3)	69.8 (3.8)	66.5 (3.2)	2.267	7.440	0.160	1.7 (1.1)
Low intensity	49.8 (3.9)	53.2 (4.0)	55.7 (4.6)	53.2 (4.8)	10.533	19.589	0.008	5.9 (1.8)
Medium intensity	76.7 (2.6)	78.7 (3.1)	75.3 (2.9)	71.6 (3.1)	0.320	36.203	0.583	−1.4 (2.5)
High intensity	77.5 (3.7)	78.6 (3.4)	78.1 (4.5)	74.4 (3.1)	0.088	24.157	0.772	0.6 (2.0)
E. Age-matched controls (<i>n</i> = 12)								
All intensities	75.6 (2.5)	76.1 (2.4)	77.0 (2.2)	—	1.767	6.873	0.211	1.4 (1.1)
Low intensity	59.9 (2.9)	64.5 (3.5)	67.9 (2.1)	—	6.911	55.909	0.023	8.0 (3.1)
Medium intensity	82.7 (3.3)	80.9 (3.3)	81.2 (2.6)	—	1.096	13.042	0.318	−1.5 (1.5)
High intensity	83.9 (2.9)	82.7 (2.4)	81.8 (3.9)	—	0.957	27.201	0.349	−2.1 (2.1)

Table 5. Caricature effects on expression recognition accuracy (percentage correct choice as anger, fear, happy, surprise, sad, or disgust) in AMD patients and age-matched controls as a function of veridical expression intensity expressed as M (SE). ^aDegrees of freedom on linear trend are AMD patients (1, 18), mild vision loss (1, 8), moderate + severe vision loss (1, 9), first-tested eyes (1, 11), age-matched controls (1,11). ^bThe p value for caricature improvement is the same as that reported for the linear trend. ^cThe 12 first-tested eyes (of the 19 analyzed for caricature effects) comprise a combination of the only eye tested from five patients and the first eye tested (which was the stronger eye) for seven patients. Bolded results highlight p values <0.005 and corresponding significant caricature improvement.

wore their best glasses for screen viewing. Free viewing was used (i.e., no chin rest or fixation) to match real-world behavior; patients were allowed to place faces in their best retinal position for viewing by moving their head sideways or up/down.

Eyes were tested monocularly (with patch over the other eye). When a patient had two eligible eyes, the stronger was tested first (to assist with ensuring patients understood instructions). For a given eye, a minimum of 328 trials (82 images \times 4 caricature levels, presented intermixed and in random order) were tested (run A). When patients were willing and fast enough to make it feasible to continue (14 eyes), the 328 trials were repeated (run B; scores averaged over the two runs). The decision to use two runs when possible was based on statistical analysis of a pilot version of the experiment using young adults shown blurred faces, which implied as many trials per patient as possible would be valuable to give error bars small enough to

test reliably for caricature effects with small numbers of eyes (e.g., as needed to support analysis of subsets of eyes in specific vision loss categories).

Before the experimental trials began, the task was explained to participants using binocular vision. All instructions were verbal. Supplement S6 details computer equipment, task instructions, and the practice phase.

Results

Table 5 shows mean expression recognition accuracy for AMD patients and age- and sex-matched controls.

AMD patients: Caricature improvements and the most effective caricature strength

We begin by analyzing AMD patients alone to include all four caricature strengths in the analysis.

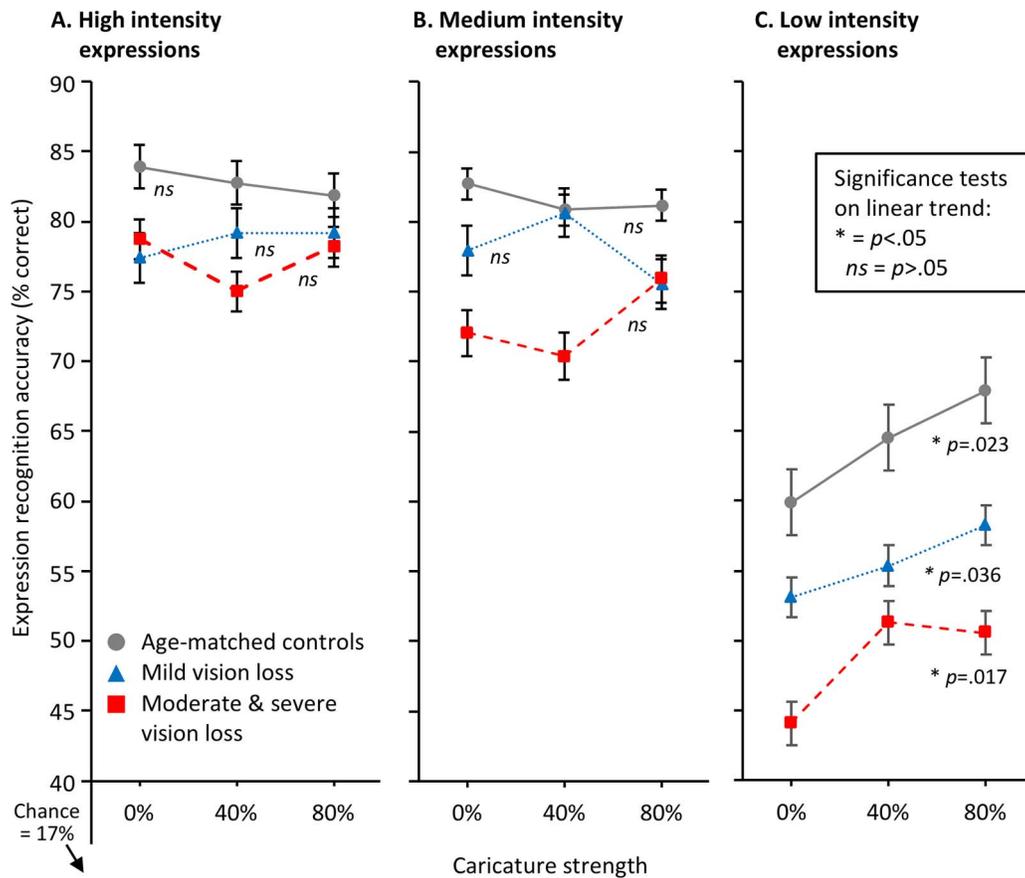


Figure 6. Caricature effects on expression recognition in AMD patients in Experiment 2. AMD-affected eyes are split into subgroups for eyes with mild vision loss eyes ($n = 9$ eyes, BCVA 6/7.5 to 6/12) and eyes with moderate-and-severe vision loss ($n = 10$ eyes, BCVA 6/19 to $< 6/360$). Data plotted up to the most effective caricature strength (80%). $p =$ significance value for linear trend across the three caricature levels shown; ns = not significant, $p > 0.05$. Error bars are the equivalent of ± 1 SEM for the repeated-measures comparison of caricature strengths (calculated separately within each vision group and intensity condition).

Note that, in this analysis, there is still a control condition, namely veridical; that is, patients are being compared to themselves to see if caricatured expressions improve performance relative to uncaricatured and, if so, what caricature strength leads to the maximum benefit in recognition accuracy.

Table 5A shows means across all 19 AMD-affected eyes (i.e., regardless of visual acuity). A two-way ANOVA (4 caricature levels \times 3 expression intensities) confirmed a main effect of intensity, $F(2, 36) = 99.202$, $MSE = 146.697$, $p < 0.001$, reflecting the fact that AMD patients, like normal vision observers in Experiment 1, found low-intensity expressions hardest to recognize. There was also a significant main effect of caricature strength, $F(3, 54) = 3.299$, $MSE = 25.59$, $p = 0.027$, consistent with AMD patients showing a caricature benefit. There was also a significant interaction between expression intensity and the linear trend on caricature, $F(1, 18) = 5.34$, $MSE = 196.587$, $p = 0.033$, showing that caricature benefits varied significantly with intensity of the veridical expression. In addition, the caricature effect also had a significant

quadratic component, $F(1, 18) = 5.265$, $MSE = 10.062$, $p = 0.034$. Table 5A shows this reflected a pattern in which patients' accuracy improved up to 80% caricature strength and then worsened with more extreme caricatures. This tendency was present for all three intensity levels, and the drop between 80% and 100% strength was significant when averaged across intensity, $t(18) = 2.36$, $p = 0.030$. Thus, the most effective caricature strength was 80%, and we conduct all further analysis examining only the first three caricature levels (i.e., 0%, 40%, and 80%).

Caricaturing up to 80% strength: Effects of vision loss severity and intensity, size of the caricature improvement, comparison to age- and sex-matched controls

Figure 6 plots results up to the maximum effective caricature strength (80%). A $3 \times 3 \times 3$ ANOVA including caricature strength (0%, 40%, 80%), vision group (matched controls, mild vision loss, moderate and severe vision loss), and veridical expression

intensity (high, medium, low) revealed no caricature \times vision group interaction, $F(4, 56) = 1.181$, $MSE = 20.299$, $p = 0.329$, and no three-way interaction, $F(8, 112) = 1.528$, $MSE = 32.756$, $p = 0.155$. These results indicate no significant changes in caricature benefit between patients and controls or within patients as a function of degree of vision loss. However, caricature strength interacted with intensity, $F(4, 112) = 3.959$, $MSE = 32.756$, $p = 0.005$, indicating a need to examine each intensity level separately.

For high-intensity expressions (Figure 6A) and with faces expanded in size for the patients, AMD patients recognized the veridical expressions very well and only slightly worse than the matched control group: patient accuracy was above 75% correct for both the mild vision loss and moderate and severe vision loss groups. At this accuracy, Experiment 1 results predict there is likely no caricature benefit, and indeed this was the case. A 3×3 ANOVA for high-intensity expressions (3 caricature strengths \times 3 vision groups) found no main effect of caricature, $F(1.471, 41.189) = 0.367$, $MSE = 33.650$, $p = 0.695$, and no interaction between caricature and vision group $F(4, 56) = 1.079$, $MSE = 24.750$, $p = 0.376$. Linear trend analysis for each vision group separately also showed no caricature benefits in any vision category (Table 5; Figure 6A).

For medium-intensity expressions (Figure 6B), veridical accuracy for mild vision loss remained good (78%) but, even with the faces expanded in size, dropped to 72% for moderate and severe vision loss. A 3×3 ANOVA (3 caricature strengths \times 3 vision groups) showed no main effect of caricature, $F(2, 56) = 0.030$, $MSE = 23.289$, $p = 0.970$, but a significant interaction between caricature and vision group, $F(4, 56) = 3.232$, $MSE = 23.289$, $p = 0.019$. Figure 6B shows that this reflects the potential emergence of a caricature benefit in which veridical recognition dropped to 72% (i.e., for moderate-to-severe vision loss). The size of this caricature benefit (i.e., 80% caricature *minus* veridical) was a 3.9% increase in accuracy. This was not significant with $n = 10$ eyes (Table 5C) although note that a benefit cannot be ruled out given that similar sized improvements were significant in Experiment 1 with a larger sample size (e.g., young adults in the low-intensity, no-blur condition showed a significant caricature benefit of 3.6%, Table 2A).

For low-intensity expressions (Figure 6C), performance for veridical expressions dropped to well below 70% correct for all vision groups, specifically, 60% for the age- and sex-matched control group, 53% for mild vision loss, and 44% for moderate-to-severe vision loss. Under this circumstance of starting from much poorer performance, caricaturing significantly improved expression recognition. A 3×3 ANOVA for low intensity confirmed a main effect of vision group with overall performance worsening with increasingly severe vision

loss, $F(2, 28) = 4.425$, $MSE = 446.439$, $p = 0.021$. More importantly, there was a main effect of caricature strength with accuracy improving across 0%, 40%, and 80% caricatures, $F(2, 56) = 9.231$, $MSE = 37.773$, $p < 0.001$, and no interaction between caricature level and vision group, $F(4, 56) = 0.586$, $MSE = 37.773$, $p = 0.674$. This latter result indicates that the size of the low-intensity expression caricature benefit did not differ between AMD patients and controls or between mild vision loss and moderate-to-severe vision loss. Analyzing each vision group independently confirmed this result. For the age- and sex-matched control group, results revealed a significant caricature improvement (linear trend across 0%, 40%, 80% caricature strength): $F(1, 11) = 6.911$, $MSE = 55.909$, $p = 0.023$, the size of which was $8.0\% \pm 3.1\%$ ($M \pm SEM$; Table 5E). For mild vision loss, results also revealed a significant caricature improvement (linear trend across 0%, 40%, 80% caricature strength): $F(1, 8) = 6.345$, $MSE = 13.384$, $p = 0.036$, the size of which was $5.1\% \pm 2.0\%$ (Table 5B). For moderate and severe vision loss, results again revealed a significant caricature improvement (linear trend across 0%, 40%, 80% caricature strength): $F(1, 9) = 8.58$, $MSE = 24.120$, $p = 0.017$, the size of which was $6.5\% \pm 2.2\%$ (Table 5C). Note the low-intensity caricaturing benefit was no weaker for moderate and severe vision loss (6.5%) than for mild vision loss (5.1%). Combining both sets of eyes together, the average caricature benefit across all acuity levels in AMD was $5.8\% \pm 1.5\%$.

A final point of note is that, for low-intensity expressions, the combination of enlarged and caricatured images improved the performance of AMD patients with mild vision loss *back to very nearly as good as normal performance for their age group*. Specifically, in real life, the normal way to see faces is veridical, and accuracy for this condition in age- and sex-matched controls was 59.9% correct; this compares to 58.2% correct in mild vision loss AMD when faces were enlarged and caricatured to the most effective strength (80% exaggeration).

Discussion

Results of Experiment 2 demonstrate that caricaturing can improve expression recognition in AMD and that, as in Experiment 1, caricature improvements emerge when they are most needed and specifically when recognition of veridical expressions dropped below approximately 72% correct. For low-intensity expressions, regardless of whether eyes had only mild vision loss or moderate-to-severe vision loss, caricaturing significantly improved patient accuracy by $5.8\% \pm 1.5\%$. This did not differ from the $8.0\% \pm 3.1\%$ benefit in the age- and sex-matched control group (nor

from $5.5\% \pm 1.7\%$ for low-intensity expressions in older adults more broadly, using all older adults from Experiment 1). Finally, the combination of expanding and caricaturing the face for patients was able to improve recognition of low-intensity expressions back to normal (uncaricatured) recognition accuracy for their age group.

A final outcome of Experiment 2 worth noting briefly concerns the effect of expression intensity on recognition of uncaricatured faces in AMD. Previous studies of AMD expression ability (Boucart et al., 2008; Johnson et al., 2017; Tejeria et al., 2002) have not considered the possibility of intensity effects nor reported intensity information for their stimuli. Here, Figure 6 shows that, when faces are enlarged (equivalent to viewing a real person from 58 cm), AMD patients achieve close to control levels of recognition accuracy for high-intensity expressions in both mild and moderate-to-severe vision loss and for medium-intensity expressions in mild vision loss patients. In contrast, deficits in expression recognition, even with enlarged faces, occur for medium-intensity faces in moderate-to-severe vision loss and for low-intensity faces even with only mild vision loss. Thus, the vision loss associated with AMD most severely impacts recognition of low-intensity expressions and, to some extent, medium-intensity expressions.

General discussion

Across both experiments, our key finding was that caricaturing can improve expression recognition in low-resolution vision. Moreover, caricature benefits on accuracy occurred when they were most needed, namely when veridical recognition was impaired. Our proof-of-concept study (Experiment 1) showed that, once performance dropped sufficiently below ceiling (to $\sim 72\%$ correct or less), caricature improvements in accuracy occurred across a wide range of conditions: for older and young adults; for low, medium, and even high intensity of original expressions; and for different resolution levels, including extremely blurred images. Our AMD patient study (Experiment 2) showed that, in patients, caricaturing again improved expression recognition when veridical recognition was poor. At the most effective caricature strength (80% exaggeration), the size of the caricature improvement was then $5.8\% \pm 1.5\%$ in AMD patients and did not differ significantly from that in age-matched controls. Importantly, caricaturing was also as effective in moderate-to-severe vision loss AMD as it was in only mild vision loss. This indicates caricaturing is of potential benefit across a wide range of AMD patients of different residual visual acuities.

Caricaturing and difficulty of the expression recognition task

Previous studies of veridical expression recognition in AMD have assessed performance using relatively easy tasks, namely simultaneous odd one out (e.g., three identical frowning images and one happy image; Tejeria et al., 2002) and a three-alternative neutral/happy/angry task (Boucart et al., 2008) and have not reported intensity information for their stimuli. In these articles, enlargement of the face image, although helpful, did not improve patients' recognition to control levels (Johnson et al., 2017). In our task requiring recognition of all six basic expressions, we extend these findings by showing that enlarging the face in patients improves recognition almost to the level of age-matched controls when expression intensity is high (Figure 6), but it fails to do so for low-intensity expressions that have the smallest physical differences from neutral.

Our caricaturing results then show that caricature benefits in low-resolution vision emerge when they are most needed, namely when the task becomes difficult enough to noticeably impair veridical recognition. This included significant caricature benefits for blurred images in normal vision observers for all veridical expression intensities (Experiment 1) and, in AMD patients, for low-intensity expressions (Experiment 2). Note that we do not wish to claim that caricaturing benefits in AMD are necessarily limited to low-intensity expressions. We found some evidence of a benefit for medium-intensity expressions under conditions in which recognition performance for these items begins to drop below ceiling; specifically, we obtained a 3.9% improvement for a combined moderate-and-severe vision loss group for whom average veridical performance had dropped to below 75%. Given that normal vision older adults in Experiment 1 showed clear caricature benefits for medium-intensity expressions when these were sufficiently blurred to more substantially impair veridical recognition (59% and 44% correct for the two blur levels), there is no reason to expect that AMD patients would not also show a significant caricature benefit for medium-intensity expressions if their overall performance was lower as would occur, for example, in a patient group restricted to all having severe vision loss or in a more demanding expression discrimination task.

Of course, in the real world, even being able to recognize the six basic expressions (Ekman, 1993) is only the bare minimum of everyday requirements for expression and emotion perception. Other important social signals sent by facial expressions can include "I'm bored with your conversation," "She's flirting" (see the Reading the Mind in the Eyes test; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), the

difference between moldy food “physical disgust” and contempt (Ekman & Friesen, 1986), or whether your grandchild is genuinely sad or merely pretending (Dawel et al., 2017). All these signals involve only small physical variations in faces, implying they are likely to be poorly perceived by patients with low-resolution vision, including AMD patients. Caricaturing offers hope of improving recognition of these types of subtle information given that both our present findings and our previous studies of simulated low vision (Dawel et al., 2019; Irons et al., 2014) show that caricaturing tends to be effective particularly when performance is initially impaired, at least as long as accuracy does not become so poor it hits floor.

Theoretical relationship between caricaturing and intensity

In the present article, our use of the term “intensity” refers to the perceived intensity of the expression in the veridical photograph. Caricaturing itself, however, also increases the perceived intensity of emotions (Calder et al., 2000). Does this mean, then, that caricaturing and natural intensity variations are equivalent? Not really. Caricaturing does increase physical intensity in the sense that it increases physical differences in an expressive image compared to a neutral expression but not necessarily in the same way that, say, a real person displaying a weaker or stronger smile does. Caricaturing can exaggerate *only the physical information that is present in a particular expression image*, but not all physical information present in a natural high-intensity face is necessarily expressed by the same person displaying a low-intensity version of the same expression. For example, the low intensity sad expression in Figure 2 contains some of the typical muscle “action units” (Ekman, Friesen, & Hager, 2002) indicating sadness, such as the downturned mouth, but it does not display the vertical forehead creases or raised inner eyebrows typically present in more intense versions of sadness.

Size of the caricature benefit and potential for additive enhancements

A key issue concerns the size of the caricature benefit. Our ~6% improvement in expression recognition accuracy is large enough to be of some practical benefit to patients. At the same time, however, 6% is only a modest improvement. Thus, rather than viewing caricaturing as a fix-all image-enhancement procedure, we see it as one of a series of additive enhancements that could be coapplied to facial images. This idea is bolstered by the fact that different enhancements derive

theoretically from independent stages of the visual processing stream; that is, mid-/high-level vision for caricaturing and low-level vision in the case of enlargement (Johnson et al., 2017; Tejeria et al., 2002) and other potential manipulations, such as increasing the contrast of certain spatial frequencies in the face (as has been applied in AMD for face identity; Peli, Goldstein, Trempe, & Arend, 1989).

We also note that our 6% caricature improvement here is for caricaturing *shape* only. Natural expressions also contain so-called “texture” information (also known as “reflectance”), which includes a number of sources of information that can potentially help improve expression recognition. These include skin coloring (e.g., fear is associated with blood drain and, thus, whitening of the skin, anger with blood inflow and, thus, reddening of the skin; Thorstenson, Elliot, Pazda, Perrett, & Xiao, 2018) and expression-relevant shadowing (e.g., which might highlight crinkles around the eyes in happy). In identity recognition, using tightly controlled stimuli with all faces photographed under the same lighting conditions, combining shape plus texture caricaturing can produce a larger benefit than caricaturing shape alone (Itz, Schweinberger, & Kaufmann, 2016). It is possible the same could occur for expression although note that a practical difficulty in more naturalistic settings is that much texture information is due to lighting conditions that are not informative about expression or emotion (e.g., redder skin due to standing in a sunset, shadows due to light coming in sideways through a window).

Technological issues in translation to patients

The long-term aim of our research program, of which this article forms one part, is to explore image-enhancement procedures in low-vision simulations and patients to determine experimentally which image manipulations actually improve behavioral performance and then to implement these manipulations on an easy-to-use patient platform so the patient can, for example, select a face from the full visual scene to track and view it enhanced (e.g., caricatured and enlarged) on a computer when video-conferencing with family or via smart glasses in real-world social interactions.

Achieving practical translation of caricature benefits to patients requires software that is able to automatically caricature faces in real time. One practical limitation of current caricaturing techniques is that they can be applied only to static images. Static images are, of course, experienced by patients (e.g., photographs on websites), and thus, improving expression recognition even of static expression images is beneficial. However, improving patients’ real-time social interactions with other people would require carica-

turing dynamic expressions. This requires technical advances within computer science. Although caricaturing itself is a solved problem (Benson & Perrett, 1991), automated assignment of enough landmark points to make an accurate expression caricature is not. With manual assignment (as also used in all previous expression caricaturing studies; Benson et al., 1999; Calder et al., 2000; Calder et al., 1997), we could accurately locate 140–230 landmark points per face. However, automatic assignment of landmark points in faces is currently restricted to a smaller number of points, e.g., 68 points in close to real time across changes in viewpoint and allowing for partial occlusion of the face such as the hand coming up to scratch the nose (Yang, He, Jia, & Patras, 2015). For identity recognition, this 68-point automatic landmark assignment procedure resulted in caricatures that are only approximately 50% as effective at improving behavioral performance as caricatures derived from hand-assigned landmarks (McKone et al., 2018). Moreover, this problem is likely to be exaggerated for expression caricaturing given that current auto-assigned locations fail to trace out many face regions relevant specifically to expression (e.g., wrinkles across the nose in disgust, exact eyebrow shape for sadness). Finally, an important additional challenge is developing methods to extract a neutral expression image of the target person from the video stream to caricature away from given that automatic expression recognition remains difficult even in constrained stimulus environments (i.e., without large changes in lighting, viewpoint, etc.; Li & Deng, 2018).

Potential for generalization to other low-vision disorders

Our results have demonstrated expression caricature benefits of roughly similar size across a wide range of conditions. This includes different *ages of observer* covering the full adult life span, from young adults (mean age 21 years) to older adults ranging in age from 65–89 years in normal vision observers and up to 93 years in AMD patients. It also includes different *forms* of low resolution in the faces, specifically spatially uniform Gaussian blur and the mix of blur, distortions, and missing parts commonly reported by AMD patients (Lane, Rohan, Sabeti, Essex, Maddess, Dawel, et al, 2018; Taylor et al., 2018). Finally, it includes different *severities* of low-resolution vision, including different levels of added blur for normal vision observers and across a wide range of levels of vision loss in AMD patients.

The good generalization of our results across different forms and severities of low-resolution vision are as expected theoretically given that caricature

benefits arise from perceptual coding of facial shape information, which occurs in mid- and/or high-level cortical visual processing areas. The good generalization across the whole adult life span also argues that, even as the brain ages, perceptual coding of expression does not significantly degrade and continues to support caricature benefits. In turn, these ideas imply that caricaturing is also likely to benefit expression recognition in vision disorders beyond those tested here, including, for example, other types of macular disease that emerge earlier in adulthood than AMD and retinitis pigmentosa.

Conclusion

We have previously shown that caricaturing *identity* is an effective way to improve identity recognition in AMD (Lane, Rohan, Sabeti, Essex, Maddess, Barnes, et al., 2018). Here, we have shown that caricaturing *expression* is an effective way to improve impaired recognition of facial emotion. Together, these findings demonstrate that face perception can be significantly improved in AMD patients by employing techniques derived theoretically from coding in mid- and high-level cortical vision. This high-level approach has the added benefit that such techniques do not depend on the specifics of retinal damage or exact visual appearance in any individual patient or any given disorder. Caricaturing has the potential to lead to practical benefits in patients a range of low-vision disorders and, indeed, even in patients without functioning eyes at all (e.g., via prosthetic implants in LGN or cortical area V1; Irons et al., 2017).

Keywords: age-related macular degeneration, facial expression, caricaturing, intensity

Acknowledgments

The Macular Disease Foundation Australia, specifically Mr. Rob Cummins, assisted with recruitment of people living with AMD. Concerning our use of some face stimuli from the NimStim database, we include this required acknowledgement: “Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development. Please contact Nim Tottenham at tott0006@tc.umn.edu for more information concerning the stimulus set.” This research was supported by Australian Research Council grants CE110001021 (www.ccd.edu.au; EM, KC, AD) and DP150100684

(EM), NHMRC Project Grant 1063458 (TM, ER, TS), Rebecca Cooper Medical Foundation Grant PG2018040 (FS), and NHMRC Project Grant 1082358 (NB). Designed experiment: JL, EM. Conducted the experiment: JL, RR, ER, FS, RE, TM, JM, JI. Analyzed/interpreted the data: JL, EM, KC, RR, MS, JI, ER, FS, RE, TM. Provided materials: AD, TG, JM, JI, EM, RE, TM, XH, NB. Wrote the article: JL, EM, KC. Proofed/revise the article: all authors.

Commercial relationships: N. Barnes, Bionic Vision Technologies (F,P).

Corresponding author: Jo Lane.

Email: jo.lane@anu.edu.au.

Address: Research School of Population Health, The Australian National University, Canberra, ACT, Australia.

Footnote

¹ Where data violated the sphericity assumption of repeated-measures ANOVA, we report Greenhouse–Geisser corrected *df*, *F*, and *MSE*.

References

- Aviezer, H., Trope, Y., & Todorov, A. (2012, November 30). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, 338(6111), 1225–1229, <https://doi.org/10.1126/science.1224313>.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241–251, <https://doi.org/10.1111/1469-7610.00715>.
- Benson, P. J., Campbell, R., Harris, T., Frank, M. G., & Tovée, M. J. (1999). Enhancing images of facial expressions. *Perception and Psychophysics*, 61(2), 259–274.
- 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(1), 105–135, <https://doi.org/10.1080/09541449108406222>.
- Boucart, M., Jean-François, D., Desprez, P., Desmettre, T., Hladiuk, K., & Oliva, A. (2008). Recognition of facial emotion in low vision: A flexible usage of facial features. *Visual Neuroscience*, 25(4), 603–609, <https://doi.org/10.1017/S0952523808080656>.
- Bunting, R., & Guymer, R. (2012). Treatment of age-related macular degeneration. *Australian Prescriber*, 35, 90–93, <https://doi.org/10.18773/austprescr.2012.038>.
- Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition*, 76, 105–146.
- Calder, A. J., Young, A. W., Rowland, D., & Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society B*, 264, 919–925.
- Dawel, A., Wong, T. Y., McMorrow, J., Ivanovici, C., He, X., Barnes, N., . . . McKone, E. (2019). Caricaturing as a general method to improve poor face recognition: Evidence from low-resolution images, other-race faces, and older adults. *Journal of Experimental Psychology: Applied*, 25(2), 256–279, <https://doi.org/10.1037/xap0000180>.
- Dawel, A., Wright, L., Irons, J., Palermo, R., O’Kearney, R., & McKone, E. (2017). Perceived emotion genuineness: Normative ratings for popular facial expression stimuli and development of perceived-as-genuine and perceived-as-fake sets. *Behavior Research Methods*, 49(4), 1539–1562.
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), 384–392, <https://doi.org/10.1037/0003-066X.48.4.384>.
- Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion*, 10(2), 159–168, <https://doi.org/10.1007/BF00992253>.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system: The manual on CD ROM*. Salt Lake City: Network Information Research Corp., Research Nexus Division.
- Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*, 85(2), 276–290, <https://doi.org/10.1037/0022-3514.85.2.276>.
- Farkas, L. G., Hreczko, T. A., & Katic, M. J. (1994). Craniofacial norms in North American Caucasians from birth (one year) to young adulthood. In L. G. Farkas (Ed.), *Anthropometry of the Head and Face* (2nd ed., pp. 241–335). New York: Raven Press.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the

- clinician. *Journal of Psychiatric Research*, 12(3), 189–198.
- Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Huggett, P., Macy, L., . . . Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, 115(2), 137–143, [https://doi.org/10.1016/S0165-0270\(02\)00006-7](https://doi.org/10.1016/S0165-0270(02)00006-7).
- Irons, J. L., Gradden, T., Zhang, A., He, X., Barnes, N., Scott, A. F., & McKone, E. (2017). Face identity recognition in simulated prosthetic vision is poorer than previously reported and can be improved by caricaturing. *Vision Research*, 137, 61–79, <https://doi.org/10.1016/j.visres.2017.06.002>.
- Irons, J. L., McKone, E., Dumbleton, R., Barnes, N., He, X., Provis, J., . . . Kwa, A. (2014). A new theoretical approach to improving face recognition in disorders of central vision: Face caricaturing. *Journal of Vision*, 14(2):12, 1–29, <https://doi.org/10.1167/14.2.12>. [PubMed] [Article]
- Itz, M. L., Schweinberger, S. R., & Kaufmann, J. M. (2016). Effects of caricaturing in shape or color on familiarity decisions for familiar and unfamiliar faces. *PLoS One*, 11(2), e0149796, <https://doi.org/10.1371/journal.pone.0149796>.
- Johnson, A. P., Woods-Fry, H., & Wittich, W. (2017). Effects of magnification on emotion perception in patients with age-related macular degeneration. *Investigative Ophthalmology and Visual Science*, 58, 2520–2526, <https://doi.org/10.1167/iovs.16-21349>.
- Kanwisher, N., & Dilks, D. D. (2013). The functional organization of the ventral visual pathway in humans. In L. Chalupa & J. Werner (Eds.), *The New Visual Neurosciences* (pp. 733–748). Cambridge, MA: The MIT Press.
- Kayaert, G., Biederman, I., Op de Beeck, H. P., & Vogels, R. (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *European Journal of Neuroscience*, 22(1), 212–224, <https://doi.org/10.1111/j.1460-9568.2005.04202.x>.
- Khandhadia, S., Cipriani, V., Yates, J. R. W., & Lotery, A. J. (2012). Age-related macular degeneration and the complement system. *Immunobiology*, 217(2), 127–146, <https://doi.org/10.1016/j.imbio.2011.07.019>.
- Kumfor, F., Irish, M., Hodges, J. R., & Piguet, O. (2013). Discrete neural correlates for the recognition of negative emotions: Insights from frontotemporal dementia. *PLoS One*, 8(6), e67457, <https://doi.org/10.1371/journal.pone.0067457>.
- Kumfor, F., Miller, L., Lah, S., Hsieh, S., Savage, S., Hodges, J. R., & Piguet, O. (2011). Are you really angry? The effect of intensity on facial emotion recognition in frontotemporal dementia. *Social Neuroscience*, 6(5–6), 502–514, <https://doi.org/10.1080/17470919.2011.620779>.
- Lane, J., Rohan, E. M. F., Sabeti, F., Essex, R. W., Maddess, T., Barnes, N., . . . McKone, E. (2018). Improving face identity perception in age-related macular degeneration via caricaturing. *Scientific Reports*, 8:15205, 1–10, <https://doi.org/10.1038/s41598-018-33543-3>.
- Lane, J., Rohan, E. M. F., Sabeti, F., Essex, R. W., Maddess, T., Dawel, A., . . . McKone, E. (2018). Impacts of impaired face perception on social interactions and quality of life in age-related macular degeneration: A qualitative study and new community resources. *PLoS One*, 13(12), e0209218, <https://doi.org/10.1371/journal.pone.0209218>.
- Leppänen, J. M., Kauppinen, P., Peltola, M. J., & Hietanen, J. K. (2007). Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. *Brain Research*, 1166, 103–109, <https://doi.org/10.1016/j.brainres.2007.06.060>.
- Li, S., & Deng, W. (2018). Deep facial expression recognition: A survey. *arXiv:1804.08348*.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces—KDEF, CD ROM: Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, Stockholm, Sweden.
- Mäkäräinen, M., Kätsyri, J., & Takala, T. (2014). Exaggerating facial expressions: A way to intensify emotion or a way to the uncanny valley? *Cognitive Computation*, 6(4), 708–721, <https://doi.org/10.1007/s12559-014-9273-0>.
- Marmor, D. J., & Marmor, M. F. (2010). Simulating vision with and without macular disease. *Archives of Ophthalmology*, 128(1), 117–125, <https://doi.org/10.1001/archophthalmol.2009.366>.
- McKone, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research*, 49(2), 268–283.
- McKone, E., Robbins, R. A., He, X., & Barnes, N. (2018). Caricaturing faces to improve identity recognition in low vision simulations: How effective is current-generation automatic assignment of landmark points? *PLoS One*, 13(10), e0204361, <https://doi.org/10.1371/journal.pone.0204361>.
- McLellan, T., Johnston, L., Dalrymple-Alford, J., & Porter, R. (2010). Sensitivity to genuine versus posed emotion specified in facial displays. *Cogni-*

- tion and Emotion, 24(8), 1277–1292, <https://doi.org/10.1080/02699930903306181>.
- Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments and Computers*, 36(4), 634–638, <https://doi.org/10.3758/BF03206544>.
- Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, 86(5), 2505–2519, <https://doi.org/10.1152/jn.2001.86.5.2505>.
- Peli, E., Goldstein, R. B., Trempe, C. L., & Arend, L. E. (1989). Image enhancement improves face recognition. In *Noninvasive Assessment of the Visual System (Vol. 7, pp. 64–67)*. Washington, DC: Optical Society of America.
- Ruffman, T., Henry, J., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, 32, 863–881, <https://doi.org/10.1016/j.neubiorev.2008.01.001>.
- Taylor, D. J., Edwards, L. A., Binns, A. M., & Crabb, D. P. (2018). Seeing it differently: Self-reported description of vision loss in dry age-related macular degeneration. *Ophthalmic and Physiological Optics*, 38, 98–105, <https://doi.org/10.1111/opo.12419>.
- Tejeria, L., Harper, R. A., Artes, P. H., & Dickinson, C. M. (2002). Face recognition in age related macular degeneration: Perceived disability, measured disability, and performance with a bioptic device. *British Journal of Ophthalmology*, 86, 1019–1026, <https://doi.org/10.1136/bjo.86.9.1019>.
- Thorstenson, C. A., Elliot, A. J., Pazda, A. D., Perrett, D. I., & Xiao, D. (2018). Emotion-color associations in the context of the face. *Emotion*, 18(7), 1032–1042, <https://doi.org/10.1037/emo0000358>.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249, <https://doi.org/10.1016/j.psychres.2008.05.006>.
- van Rheede, J. J., Wilson, I. R., Qian, R. I., Downes, S. M., Kennard, C., & Hicks, S. L. (2015). Improving mobility performance in low vision with a distance-based representation of the visual scene. *Investigative Ophthalmology and Visual Science*, 56, 4802–4809, <https://doi.org/10.1167/iovs.14-16311>.
- Wegrzyn, M., Riehle, M., Labudda, K., Woermann, F., Baumgartner, F., Pollmann, S., . . . Kissler, J. (2015). Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. *Cortex*, 69, 131–140, <https://doi.org/10.1016/j.cortex.2015.05.003>.
- World Health Organization. (2015). *The International Statistical Classification of Diseases and Related Health Problems (ICD-10)*. Geneva: World Health Organization.
- Yang, H., He, X., Jia, X., & Patras, I. (2015). Robust face alignment under occlusion via regional predictive power estimation. *IEEE Transactions on Image Processing*, 24(8), 2393–2403, <https://doi.org/10.1109/TIP.2015.2421438>.