

Caricature generalization benefits for faces learned with enhanced idiosyncratic shape or texture

Marlena L. Itz¹ · Stefan R. Schweinberger^{1,2} · Jürgen M. Kaufmann^{1,2}

Published online: 7 October 2016
© Psychonomic Society, Inc. 2016

Abstract Recent findings show benefits for learning and subsequent recognition of faces caricatured in shape or texture, but there is little evidence on whether this caricature learning advantage generalizes to recognition of veridical counterparts at test. Moreover, it has been reported that there is a relatively higher contribution of texture information, at the expense of shape information, for familiar compared to unfamiliar face recognition. The aim of this study was to examine whether veridical faces are recognized better when they were learned as caricatures compared to when they were learned as veridicals—what we call a *caricature generalization benefit*. Photorealistic facial stimuli derived from a 3-D camera system were caricatured selectively in either shape or texture by 50 %. Faces were learned across different images either as veridicals, shape caricatures, or texture caricatures. At test, all learned and novel faces were presented as previously unseen frontal veridicals, and participants performed an old–new task. We assessed accuracies, reaction times, and face-sensitive event-related potentials (ERPs). Faces learned as caricatures were recognized more accurately than faces learned as veridicals. At learning, N250 and LPC were largest for shape caricatures, suggesting encoding advantages of distinctive facial shape. At test, LPC was largest for faces that had been learned as texture caricatures, indicating the importance of texture for familiar face recognition. Overall, our findings demonstrate that

caricature learning advantages can generalize to and, importantly, *improve* recognition of veridical versions of faces.

Keywords Face learning · Shape · Texture · Caricaturing · Encoding · ERPs

It has been shown consistently that distinctive faces are easier to remember than prototypical faces (Hancock, Burton, & Bruce, 1996; Sommer, Heinz, Leuthold, Matt, & Schweinberger, 1995; Vokey & Read, 1992). In line with a norm-based coding model within a multidimensional face space framework (Valentine, 1991; Valentine, Lewis, & Hills, 2014), distinctive faces, compared to prototypical faces, lie in more peripheral areas of a face space. Thus, the recognition advantage of distinctive faces is thought to stem from their lower likelihood to be mistaken for other faces.¹ One way to explore this effect is by using caricatures (i.e., faces whose idiosyncratic characteristics that deviate from a norm have been exaggerated; Perkins, 1975). Computerized caricatures for line drawings of faces were pioneered by Brennan (1985), but high-quality photorealistic caricatures were invented by Benson and Perrett (1991). Today, digital caricaturing can be applied to facial shape or texture (e.g., Itz, Schweinberger, Schulz, & Kaufmann, 2014). In terms of morphable models, facial shape corresponds to *warping*, whereas facial texture corresponds to *fading* (Beale & Keil, 1995). Accordingly, shape refers to the overall form of a face as well as the form of individual facial features and their second-order configuration (e.g., interocular distance), whereas texture refers to RGB color space. Several face-learning studies compared caricatures with veridical faces in conditions in which individual faces were presented in the same format at

✉ Marlena L. Itz
marlena.itz@uni-jena.de

¹ Department of General Psychology and Cognitive Neuroscience, Institute of Psychology, Friedrich Schiller University of Jena, Am Steiger 3, Haus 1, 07743 Jena, Germany

² DFG Research Unit Person Perception, Friedrich Schiller University of Jena, Am Steiger 3, Haus 1, 07743 Jena, Germany

¹ Please see Burton and Vokey (1998) for a discussion on the distribution of faces within a multidimensional face space model.

both learning and test. In those conditions, consistent recognition performance advantages and modulation of face-sensitive event-related potentials (ERPs) have been found for shape caricatures (Itz et al., 2014; Kaufmann, Schulz, & Schweinberger, 2013; Schulz, Kaufmann, Kurt, & Schweinberger, 2012; Schulz, Kaufmann, Walther, & Schweinberger, 2012). Recently, similar and even larger caricature benefits were found for faces caricatured in texture (Itz et al., 2014). Note, however, that because faces that were learned as caricatures were also presented as caricatures at test, it remains unclear whether caricaturing benefits at learning would generalize to superior recognition of veridicals at test. Please note the difference to the “superportrait” idea (Rhodes, 1996), according to which faces learned as veridicals are recognized better when *tested* as caricatures. Our aim in this study was to examine whether learning faces as caricatures, versus learning them as veridicals, improves subsequent recognition of their veridical counterparts. If so, this would have important theoretical implications for models of face learning as well as practical implications (e.g., for designing training programs for people with poor face-recognition skills).

Research on such a *caricature generalization benefit* is virtually absent for photorealistic faces. Most studies that have been published so far used incomplete facial stimuli with low degrees of realism. Note also that all studies to date on this effect used shape caricatures only. Stevenage (1995) found fewer errors for identification of veridical facial photographs that had been learned as line-drawn caricatures compared to veridical facial photographs that had been learned as line-drawn veridicals. These findings were, however, restricted to short periods of learning and vanished with unlimited learning. Later studies using 3-D face scans devoid of texture information found caricature generalization benefits (termed “reverse caricature effect” in those studies) as indicated by higher recognition accuracy (Deffenbacher, Johanson, Vetter, & O’Toole, 2000; Rodriguez, Bortfeld, Rudomin, Hernandez, & Gutierrez-Osuna, 2009). Two more recent studies, however, failed to find replicate those findings (Kaufmann & Schweinberger, 2012; Rodriguez & Gutierrez-Osuna, 2011), which may have been due to usage of either very mild levels of caricaturing (30 %; Kaufmann & Schweinberger, 2012) or artificial 3-D faces reconstructed from frontal photographs of faces (Rodriguez & Gutierrez-Osuna, 2011).

Neural correlates from ERP studies shed further light on encoding advantages for distinctive and caricatured facial stimuli. In the first study investigating ERP modulation by caricaturing, Kaufmann and Schweinberger (2008) found evidence for an encoding advantage for photorealistic shape caricatures of unfamiliar faces. In that study, participants performed a simple familiarity classification task on preexperimentally familiar and unfamiliar faces. The authors

found larger occipitotemporal N170 and N250 for *unfamiliar* shape caricatures compared to veridicals, whereas this pattern of results was absent for personally familiar faces. This led the authors to hypothesize that shape caricaturing enhances learning of unfamiliar faces in particular, but not recognition of preexperimentally familiar faces. Later face-learning studies using shape caricatures replicated this pattern for N170 (Kaufmann & Schweinberger, 2012; Schulz, Kaufmann, Kurt, et al., 2012) and N250 (Itz et al., 2014; Kaufmann et al., 2013; Kaufmann & Schweinberger, 2012; Schulz, Kaufmann, Kurt, et al., 2012; Schulz, Kaufmann, Walther, et al., 2012). Moreover, the finding of larger N250 amplitudes was extended to texture caricatures (Itz et al., 2014). Whereas N170 is typically associated with preidentity “structural encoding” of faces (Eimer, 2000b), larger (i.e., more negative) N250 is attributed to the processing of familiarity (Kaufmann, Schweinberger, & Burton, 2009; Schweinberger & Neumann, 2016; Tanaka, Curran, Porterfield, & Collins, 2006). The observations of larger N250 in face-learning paradigms for distinctive (e.g., caricatures, naturally distinctive, and other-race faces) unfamiliar compared to more typical unfamiliar faces (Schulz, Kaufmann, Kurt, et al., 2012; Wiese, Kaufmann, & Schweinberger, 2014; see also Sommer et al., 1995, for related findings) suggest that N250 is associated not only with familiarity but also with the encoding of distinctive facial information.

A further ERP component of interest, the occipitotemporal P200, or P2, has been associated with facial typicality (Latinus & Taylor, 2006; Schulz, Kaufmann, Walther, et al., 2012). There are consistent and strong findings of smaller P200 for shape caricatures (Itz et al., 2014; Kaufmann & Schweinberger, 2012; Schulz, Kaufmann, Kurt, et al., 2012; Schulz, Kaufmann, Walther, et al., 2012), as well as for other-race faces (e.g., Stahl, Wiese, & Schweinberger, 2008; Wiese, Kaufmann, et al., 2014). Significant effects of smaller P200 have also been reported for texture caricatures (Itz et al., 2014) and naturally distinctive faces (Schulz, Kaufmann, Kurt, et al., 2012). These findings are in line with postulations that P200 is modulated by typicality, with smaller amplitudes for “untypical” compared to typical faces, especially in terms of shape (Schulz, Kaufmann, Walther, et al., 2012).

Finally, a late-positive component (LPC) found at central and parietal electrode sites is larger for preexperimentally familiar (Eimer, 2000a) and experimentally learned (Kaufmann et al., 2009) compared to unfamiliar faces and has been associated with postperceptual processing of familiarity and semantic information (Schweinberger, Pfütze, & Sommer, 1995). Moreover, LPC is consistently larger for caricatured compared to veridical facial stimuli (Itz et al., 2014; Kaufmann & Schweinberger, 2012; Schulz, Kaufmann, Kurt, et al., 2012; Schulz, Kaufmann, Walther, et al., 2012).

Thus, larger LPC for caricatured compared to noncaricatured stimuli may reflect better representations and/or more efficient or deeper semantic processing of caricatures.

Converging evidence suggests that while shape is important for the initial encoding of faces, texture information, at the expense of shape, becomes more important for face recognition with increasing familiarity (see also Burton, Schweinberger, Jenkins, & Kaufmann, 2015). For instance, familiar face recognition is surprisingly robust to shape normalization (Burton, Jenkins, Hancock, & White, 2005; Russell & Sinha, 2007), as well as to massive facial shape distortion by means of image stretching (Bindemann, Burton, Leuthold, & Schweinberger, 2008; Hole, George, Eaves, & Rasek 2002). Note also that Itz et al. (2014) found largest performance benefits of texture caricaturing for recognition of learned faces and largest benefits of shape caricaturing for correct rejections of novel faces.

We tested whether encoding benefits for caricatured faces generalize to veridical test faces. To the best of our knowledge, this is the first study to investigate directly such a *caricature generalization benefit* and its underlying neural correlates for photorealistic shape or texture caricatures. Faces were learned across different images either as veridicals, shape caricatures, or texture caricatures. At test, all faces (i.e., learned and novel ones) were shown as previously unseen frontal veridicals. Thus, different images were used at learning and test for all conditions. Overall, we expected higher accuracies and shorter reaction times for faces that had been learned as caricatures compared to those that had been learned as veridicals. Moreover, we expected smaller P200 for shape caricatures and larger N250 and LPC for both types of caricatures at learning. In the ERP data at test, we expected larger N250 and LPC for learned overall compared to novel faces. Moreover, we were interested in whether ERPs at test would be modulated by the factor learning condition.

Material and method

Participants

We collected data from 36 participants (eight males) ages 18 to 32 years (mean age = 21.5 years), two of whom were left-handed (Oldfield, 1971). All participants reported to have either normal or corrected-to-normal vision. Participants were compensated monetarily or with course credit for their participation in the study. Data from one additional participant were excluded due to insufficient

EEG data quality. Participants provided written informed consent, and this study was carried out in accordance with the Declaration of Helsinki and was approved by the Ethical Commission of the Faculty of Social and Behavioural Sciences at the Friedrich Schiller University of Jena.

Stimuli

Experimental stimuli consisted of 90 unfamiliar, Caucasian faces showing neutral facial expressions generated from a 3-D camera system, used in Itz et al. (2014). Faces were captured using DI3Dcapture and then standardized and caricatured in DI3Dview (Dimensional Imaging, Glasgow, UK, Version 5.2.1). Faces were standardized to meshes provided by Dimensional Imaging containing 1,200 vertices: Using the shape transfer plugin in DI3Dview, we manually placed 44 reference points on the face in order to allocate the face to the correct positions on the standardized meshes that were then reshaped to take on the shape of each individual face. Texture maps were then created by allocating pixel information to the correct quadrants of the shape meshes. Faces were then aligned using the world alignment plugin. We then generated a female average and a male average face with the generate average OBJ plug-in. By using the morph plug-in, caricatures were generated by extrapolating each individual face with respect to the gender-matched average in either shape or texture while holding the other dimension constant, such that deviations of the individual faces from the average were enhanced by 50 %. For each identity and face type condition (veridical, shape caricature, and texture caricature), seven images were then generated and saved as bitmaps in Photoshop (CS4): one frontal and six additional viewing angles as described in Itz et al. (2014), that is, 10° downward, 10° downward and 15° to the left, 10° downward and 15° to the right, 10° upward, 10° upward and 15° to the left, and 10° upward and 15° to the right. Note that within each condition the seven images were derived from the same OBJ file for each identity. Therefore, all seven images were from the same snapshot and differed only in the rendered angle. Luminance of the final images was not altered. Faces were presented with EPrime (Version 2.0) on a black background (RGB: 0) in the center of a 16-in. color monitor with a screen resolution of 1,280 × 1,024 pixels. Final image size was 640 × 640 pixels. Using a chin rest, viewing distance was 90 cm, and face size was approximately 10 × 8 cm, for an approximate viewing angle of 6° × 5°. For examples of our stimuli, please see Fig. 1.

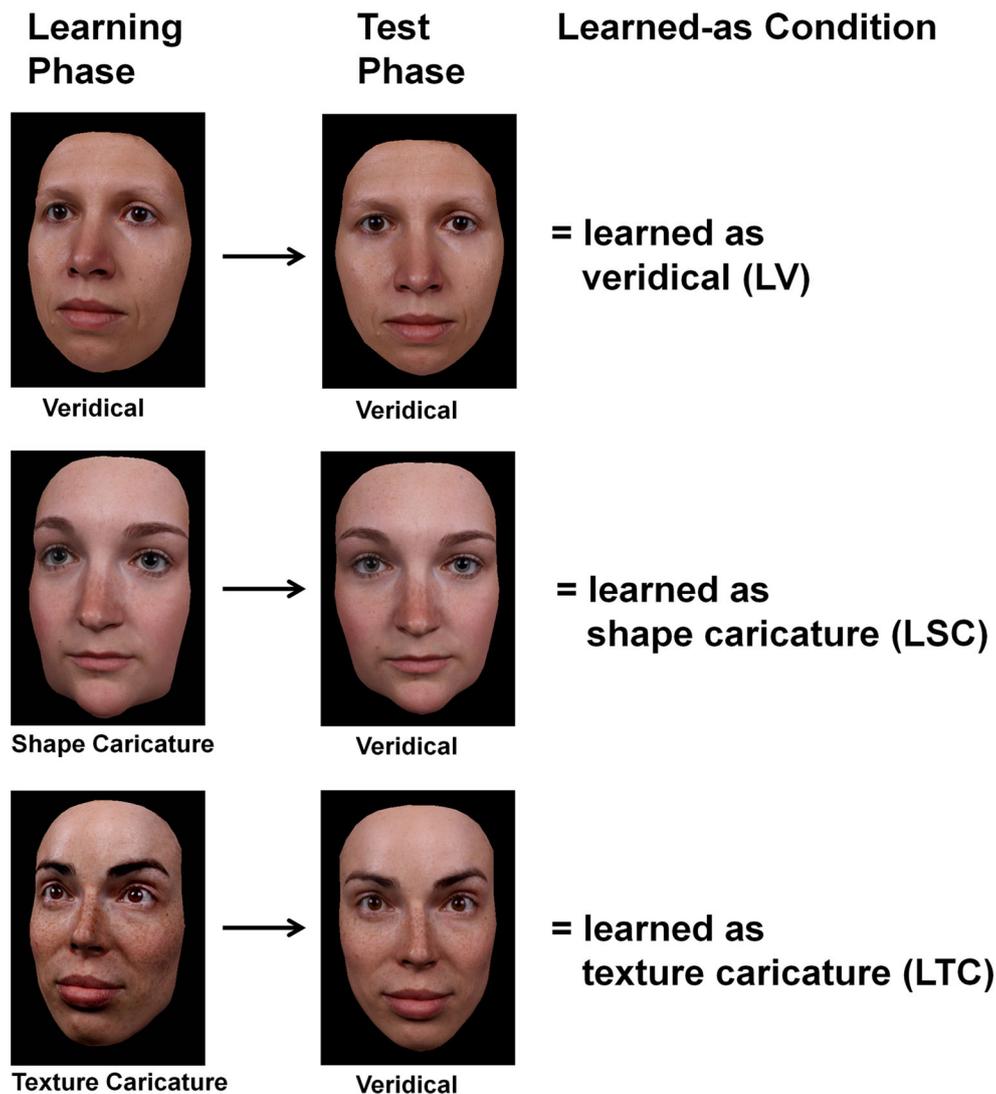


Fig. 1 Stimulus examples. Faces were caricatured in either shape or texture by 50 %. Note that faces were learned across six different viewing angles (see Fig. 1 in Itz et al., 2014) and at test all faces were presented as previously unseen veridicals

Procedure

The experiment consisted of three blocks, each with a learning phase and a subsequent test phase. Each block began with a learning phase in which 15 identities (five veridicals, five shape caricatures, and five texture caricatures) were presented at random as triplets. Each triplet was presented two times in total, and after each triplet participants were instructed to rate on a 6-point Likert scale for perceived attractiveness (after the first triplet) and distinctiveness (after the second triplet, i.e., how likely a presented face can be detected in a crowd; see Valentine & Bruce, 1986) of each identity. A learning trial consisted of the following: a fixation cross for 500 ms, a face for 3,000 ms, another image of the same face for

3,000 ms, an additional image of the same face for 3,000 ms (see below for details on the images used), the 6-point Likert scale rating (unlimited time window), and finally a blank screen for 500 ms. Faces were thus learned across different images and shown with exemplars 10° downward; 10° downward and 15° to the left; and 10° upward and 15° to the right during the first presentation. During the second presentation, faces were presented with exemplars 10° downward and 15° to the right; 10° upward; and 10° upward and 15° to the left. The ratings were not factored in to the data analyses and served simply to intensify learning and make the learning phases more interesting for the participants. Within each triplet, the exact order of the three faces was random. The order of identity presentation was also random. In total, 48

identities were learned, three of which were practice items in the first block.

After each learning phase, a test phase followed. All faces (learned and novel) were presented as frontal veridicals. A trial consisted of a fixation cross for 500 ms, followed by the frontal veridical face (presented until keypress response or for a maximum of 2,000 ms). An additional screen followed for 1,200 seconds, which was either blank or on which “Too slow!” (“Zu langsam!” in German) was presented in case participants did not answer within the 2,000-ms stimulus presentation time window. The task was to indicate as accurately and as quickly as possible whether the presented face was learned or novel. The task was speeded so that reaction times could be analyzed, in addition to accuracies. Each identity was presented a total of three times (but not as a triplet). Key assignment was counterbalanced across participants. Also, assignment of each specific identity to learned as conditions (i.e., veridical, shape caricature, or texture caricature) as well as familiarity (novel vs. learned) were counterbalanced across participants.

Six additional identities (three learned and three novel) were used as practice stimuli and were presented within the first block of the experiment. During the first test phase, practice items were presented at the beginning, and feedback was given for these items to ensure that participants had understood the task.

The experiment lasted around 40 minutes, including ample self-timed breaks between each block as well as between each learning and test phase.

Behavioral data recordings

We analyzed accuracies as well as reaction times (RTs). Note that we analyzed RTs for correct responses only. Responses not given within a time window of 200 to 2,000-ms poststimulus onset were not included in the analyses.

Electrophysiological recording and analyses

Data were recorded in an electrically shielded room. Electroencephalographic (EEG) data were recorded using sintered Ag/AgCl electrodes attached to an EasyCap electrode cap (Herrsching-Breitbrunn, Germany). Electrodes were arranged in accordance to the extended 10–20 system at scalp positions Fz, Cz, Pz, Iz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, FT9, FT10, P9, P10, PO9, PO10, F9, F10, F9', F10', TP9, and TP10. Cz composed a reference and AFz, a forehead electrode, composed ground. From electrodes (F9' and F10') placed on the

outer canthi of both eyes, horizontal electrooculogram (EOG) signals were recorded; from electrodes places above and below the left eye, vertical EOG signals were recorded. EEG data were amplified with SynAmps amplifiers (NeuroScan Labs, Sterling, VA). Signals were recorded using AC (0.05–100 Hz, -6 dB attenuation, 12 db/octave) and a sampling rate of 500 Hz. Impedances were kept below 10 k Ω .

Ocular artifacts were corrected offline automatically using BESA 5.1 (see Berg & Scherg, 1994). We generated epochs between -200-ms prestimulus onset and 1,100-ms poststimulus onset, with -200–0 ms serving as a baseline. We excluded trials contaminated with nonocular artifacts (amplitude threshold of 120 μ V, with a gradient criterion of 75 μ V) from further analyses. We averaged ERPs across blocks for learning and test conditions. Note that to avoid possible adaptation effects (Harris & Nakayama, 2007; Maurer, Rossion, & McCandliss, 2008), for ERPs in the learning phase, only amplitudes for images at the first presentation position of each triplet were analyzed (see the Procedure section). For the test phase, only trials with correct responses (i.e., hits and correct rejections) were analyzed. Averaged ERPs were then low-pass filtered at 20 Hz (zero phase shift; 12 db/octave) and recalculated to average reference. Vertical and horizontal EOG electrodes were excluded.

ERPs were calculated relative to a 200-ms prestimulus baseline. For the learning phase, mean amplitudes for the following time windows were analyzed: P100 (105–145 ms), N170 (160–200 ms), P200 (240–280 ms), N250 (280–350 ms). For the test phase, the following time windows were analyzed: P100 (95–135 ms), N170 (145–185 ms), P200 (210–250 ms), N250 (250–320 ms). For both learning and test phases, two central late-positive components were also analyzed: LPCa (500–650 ms), and LPCb (650–800 ms). Time intervals for P100, N170, and P200 were chosen based on distinct peaks identified in the grand mean averages across all conditions at learning (125, 181, and 261 ms) and test (114, 166, and 229 ms). Time-windows for N250, LPCa, and LPCb were chosen based on both visual inspection of the means and previous research on these components. Note that the peaks of each component for each individual participant lay within the chosen time windows. We chose to split the LPC into two components because of a previous study examining shape and texture caricaturing and face learning that found different results for the two time windows (Itz et al., 2014). Here, visual inspection of the means also suggested a different pattern of results for the earlier

Table 1 Behavioral measures

| | Accuracies | Mean RTs |
|-----------------|---------------------|---------------|
| Learned as V | 0.618 (0.561–0.674) | 906 (863–948) |
| Learned as SC | 0.697 (0.650–0.743) | 878 (842–914) |
| Learned as TC | 0.711 (0.668–0.754) | 883 (842–924) |
| Learned Overall | 0.675 (0.635–0.715) | 889 (851–927) |
| Novel | 0.898 (0.879–0.918) | 895 (852–937) |

Note. Mean accuracies (proportion correct) and reaction times (RTs; in milliseconds) for all conditions. 95 % confidence intervals are in parentheses. Learned as V = learned as veridicals; Learned as SC = learned as shape caricatures; Learned as TC = learned as texture caricatures

and the later parts of the LPC. P100 was quantified at O1/O2; N170 and P200 were quantified at P7/P8, P9/P10, and PO9/PO10; N250 was quantified at P7/P8, P9/P10, PO9/PO10, and O1/O2; and LPCa and LPCb were quantified at C3, Cz, and C4. The average numbers of trials per condition for the learning phase were 29.0 (veridicals), 29.2 (shape caricatures), and 29.0 (texture caricatures); and for the test phase were 26.8 (learned as veridical), 29.9 (learned as shape caricatures), 31.3 (learned as texture caricatures), and 117.5 (novel).

Results

Note that epsilon corrections heterogeneity of covariances were conducted throughout (Huynh & Feldt, 1976). In

cases of three post hoc pairwise comparisons, the alpha level was Bonferroni adjusted to .017 (Abdi, 2007).

Behavioral data

Accuracies and mean reaction times were analyzed with one-way repeated-measures ANOVAs with the factor learned-as condition (learned as veridicals [LV], learned as shape caricatures [LSC], & learned as texture caricatures [LTC]). Post hoc comparisons were analyzed with pairwise *t* tests. Moreover, pairwise *t* tests were used to analyze the comparison of learned overall (averaged across the three aforementioned learned as conditions) versus novel.

Accuracies

The one-way ANOVA revealed a main effect of learned-as condition, $F(2, 70) = 8.60, p = .001, \eta_p^2 = .197, \epsilon_{HF} = .797$. Pairwise *t* tests yielded higher accuracies for faces learned as either shape (LSC), $t(35) = 3.82, p = .001$, or texture (LTC), $t(35) = 3.09, p = .004$, caricatures when compared to those learned as veridicals (LV), with no difference between LSC and LTC, $t(35) = 0.69, p = .493$ (see Table 1 and Fig. 2). Moreover, accuracies were higher for novel compared to learned faces overall, $t(35) = 8.93, p < .001$, reflecting a common tendency to respond conservatively and rather miss a learned face than make a false alarm to a novel face.

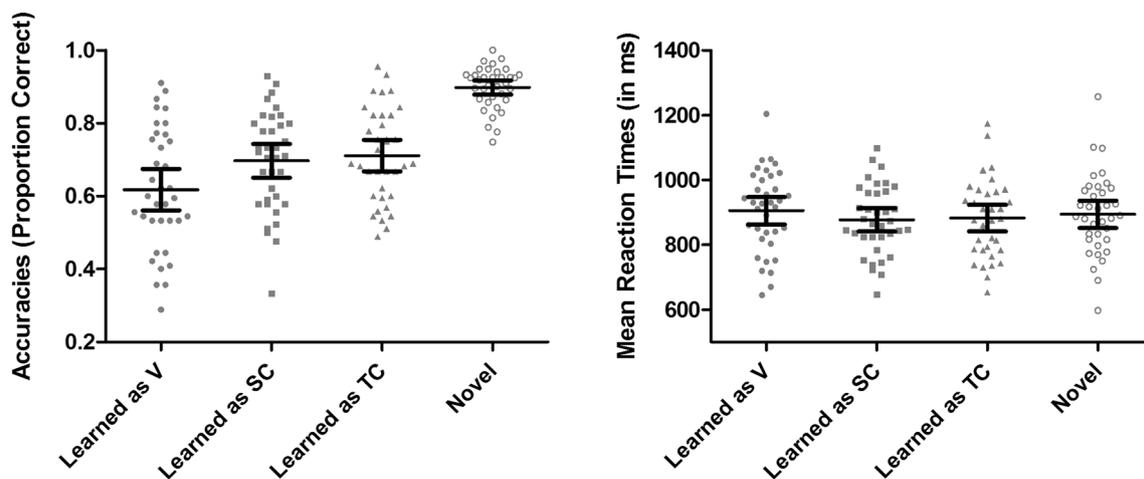


Fig. 2 Mean accuracies (proportion correct, left) and mean reaction times (in ms, right) for learned and novel faces. Note that all faces (learned and novel) were presented as frontal veridicals in the test phase. Learned as V = learned as veridical; Learned as SC = learned as shape caricature;

Learned as TC = Learned as texture caricature; Novel = faces that were not presented at learning phase. Vertical bars represent 95 % confidence intervals (CIs)

Reaction times

For mean reaction times, there was a main effect of learned-as condition, $F(1, 35) = 3.65$, $p = .031$, $\eta_p^2 = .094$. Reaction times were shorter for LSC, $t(35) = 2.45$, $p = .019$, compared to LV, although this comparison failed to reach the adjusted significance level of .017. Moreover, reaction times were numerically shorter for LTC compared to LV, $t(35) = 1.84$, $p = .075$. There was no difference between LSC and LTC, $t(35) = 0.57$, $p = .574$ (see Table 1 and Fig. 2). Moreover, mean reaction times did not differ between learned overall compared to novel faces, $t(35) = 0.49$, $p = .629$.

In summary, veridical faces that had been learned as either shape or texture caricatures were recognized better than faces that had been learned as veridicals. This is indicated by significantly higher accuracies for both types of caricatures compared to veridicals, as well as by similar trends for faster mean reaction times.

Electrophysiological data

ERP data were analyzed with ANOVAs with repeated measurements on the factor learned-as condition (i.e.,

learned as veridicals [LV], vs. learned as shape caricatures [LSC], vs. learned as texture caricatures [LTC]). For P100 the additional factor of hemisphere was included; for N170, P200, N250 the additional factors of site and hemisphere were included; and for LPCa and LPCb the additional factor of laterality was included. To investigate effects of learned overall compared to novel faces, ANOVAs without the factor of learned-as condition but with the factor familiarity (learned overall vs. novel) were used.

Learning phase

In the learning phase there were no significant effects, interactions, or trends involving face type for ERPs, P100 and N170. For P200 there was a trend for the main effect of face type, $F(2, 70) = 2.53$, $p = .087$, $\eta_p^2 = .067$, due to smaller amplitudes for shape caricatures (SC; see Fig. 3 & Table 2).

The first significant ERP effect of face type at learning emerged in the N250 time window, $F(2, 70) = 4.39$, $p = .016$, $\eta_p^2 = .111$, due to larger N250 for SC compared to veridicals (V), $F(1, 35) = 7.55$, $p = .009$,

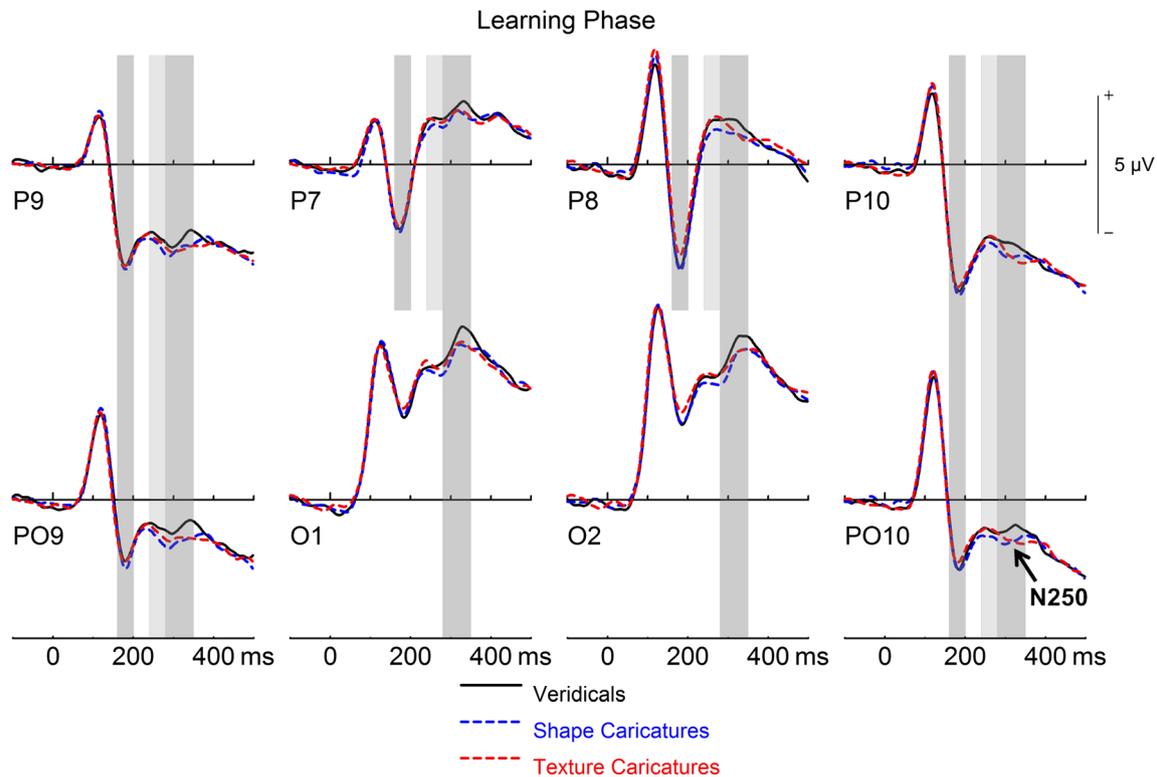


Fig. 3 Learning phase ERPs at occipitotemporal electrodes. Gray-shaded areas represent time windows of interest. N170 (160–200 ms) and P200 (240–280 ms) were analyzed at P7/P8, P9/P10, and PO9/

PO10; N250 (280–350 ms) was analyzed at all electrodes depicted in this figure

Table 2 Descriptive EEG data for significant main effects and trends of face type (Learning Phase), learned-as condition (Test Phase), and familiarity (Test Phase)

| Phase | Condition | | | |
|----------|-----------|------------------------|------------------------|------------------------|
| | ERP | V | SC | TC |
| Learning | P200 | -0.706 (-1.277--0.135) | -1.011 (-1.622--0.399) | -0.750 (-1.392--0.107) |
| | N250* | 0.906 (0.268–1.543) | 0.470 (-0.182–1.122) | 0.553 (-0.133–1.240) |
| | LPCa | 4.370 (3.697–5.044) | 5.005 (4.356–5.654) | 4.768 (3.994–5.541) |
| | LPCb* | 4.491 (3.869–5.113) | 5.345 (4.682–6.008) | 5.056 (4.317–5.795) |
| Test | | LV | LSC | LTC |
| | P100 | 4.944 (3.877–6.011) | 4.572 (3.436–5.709) | 4.710 (3.680–5.740) |
| | LPCb* | 5.742 (4.895–6.589) | 6.179 (5.200–7.159) | 6.372 (5.428–7.317) |
| Test | | Learned Overall | Novel | |
| | N250* | -1.459 (-2.312--0.606) | -1.247 (-2.076--0.417) | |
| | LPCa* | 6.118 (5.254–6.982) | 4.547 (3.735–5.360) | |
| | LPCb* | 6.098 (5.219–6.977) | 4.626 (3.838–5.414) | |

Note. Mean amplitudes and corresponding 95 % confidence intervals (in μV) for trends and significant effects of face type (learning phase), learned-as (test phase) conditions, and familiarity (learned vs. novel; test phase). V = veridicals; SC = shape caricatures; TC = texture caricatures; LV = learned as veridicals; LSC = learned as shape caricatures; LTC = learned as texture caricatures

*indicates significance of main effect with α level of 0.05

$\eta_p^2 = .177$, and a tendency for larger N250 for texture caricatures (TC) compared to V, $F(1, 35) = 4.86$, $p = .034$, $\eta_p^2 = .122$ (see Fig. 3 & Table 2). There was no difference between SC and TC, $F(1, 35)$

$= .310$, $p = .581$, $\eta_p^2 = .009$. Moreover, there were no interactions of face type with site and/or hemisphere.

For the LPCa, there was a trend for a main effect of face type, $F(2, 70) = 2.68$, $p = .075$, $\eta_p^2 = .071$, which became

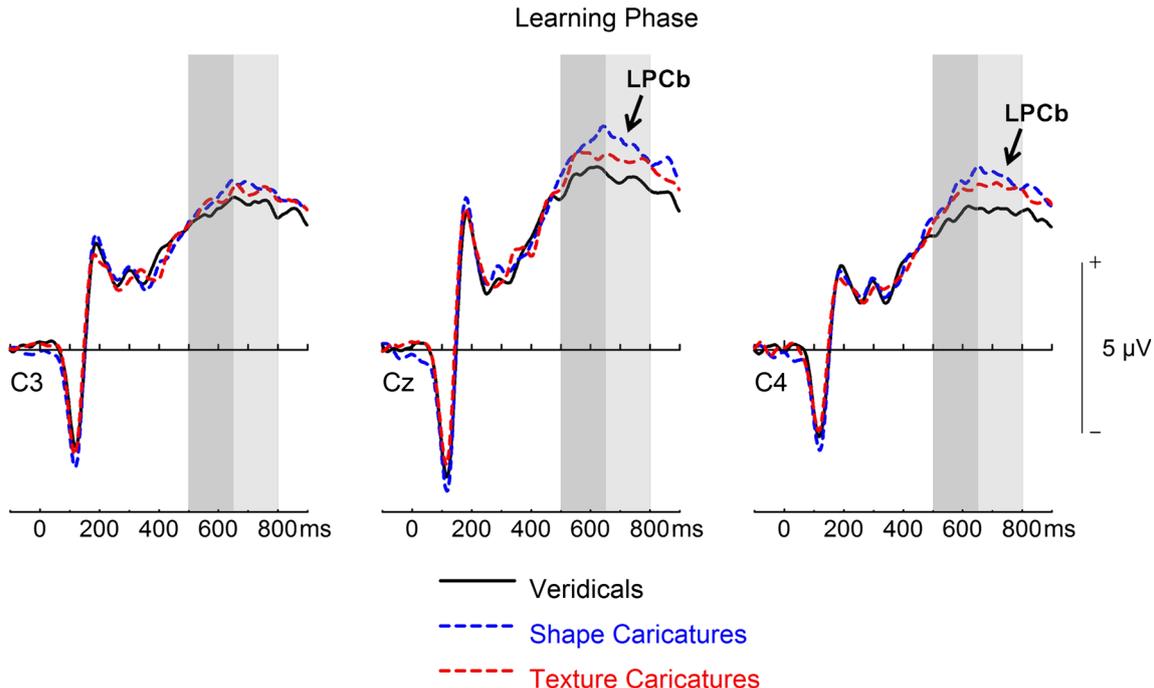


Fig. 4 Later ERPs in the learning phase at central electrodes. Gray-shaded areas represent time windows of interest: LPCa (500–650 ms) and LPCb (650–800 ms)

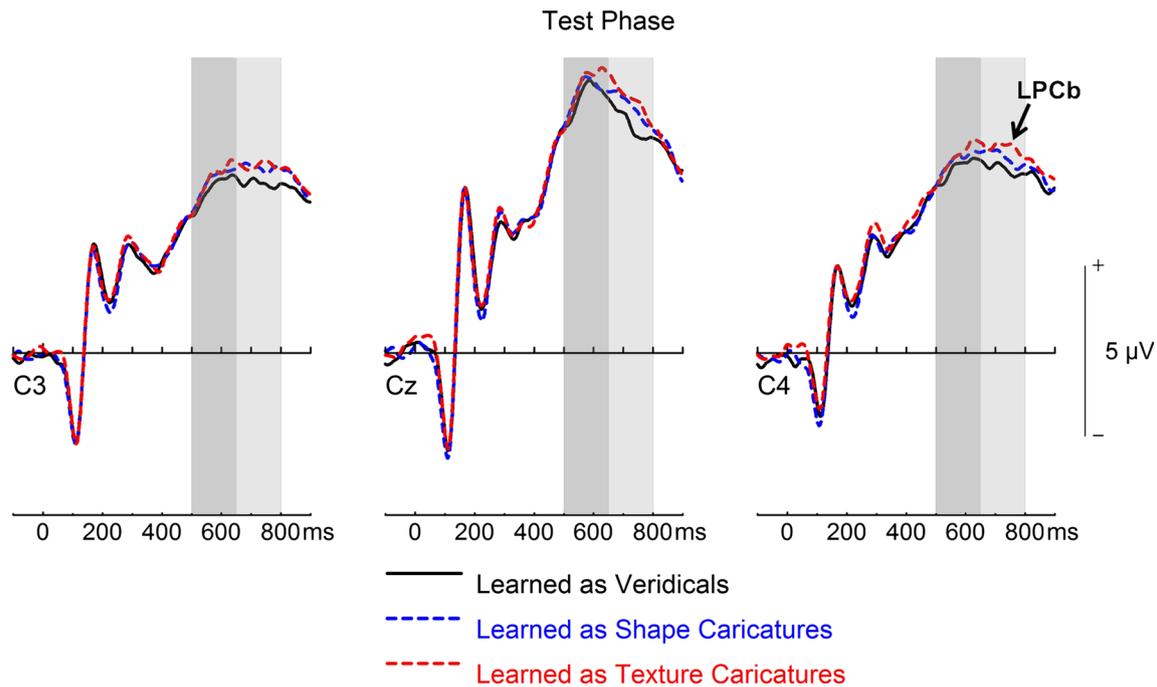


Fig. 5 Later ERPs in the test phase at central electrodes. Gray-shaded areas depict time windows of interest: LPCa (500–650 ms) and LPCb (650–800 ms)

significant for the later LPCb, $F(2, 70) = 4.52$, $p = .014$, $\eta_p^2 = .114$, due to larger amplitudes for SC compared to V, $F(1, 30) = 13.22$, $p = .001$, $\eta_p^2 = .274$ (see Fig. 4 & Table 2). The comparison between TC and V pointed into the same direction, but only yielded a trend, $F(1, 35) = 3.06$, $p = .089$, $\eta_p^2 = .080$. There was no difference between SC and TC amplitudes, $F(1, 35) = 0.92$, $p = .345$, $\eta_p^2 = .026$. For LPCb, there was also a trend for the two-way interaction laterality by face type, $F(4, 140) = 2.15$, $p = .077$, $\eta_p^2 = .058$, which reflected that these effects appeared to be slightly more prominent over the midline and right hemisphere, but smaller over the left hemisphere.

Overall, at learning, faces that were presented as shape caricatures elicited larger N250 and LPCb than did faces that were presented as veridicals, suggesting an encoding advantage for distinctive facial shape.

Test phase

For the ERP components N170, P200, N250, and LPCa there were no significant effects, interactions, or trends involving the factor learned-as condition. For P100 there was only a trend for learned-as condition, $F(2, 70) = 2.44$, $p = .095$, $\eta_p^2 = .065$. Interestingly, a significant main effect for learned-as condition was found for LPCb, $F(2, 70) = 3.40$, $p = .039$, $\eta_p^2 = .089$, due to

larger amplitudes for LTC compared to LV, $F(1, 30) = 6.01$, $p = .019$, $\eta_p^2 = .147$ (see Fig. 5 & Table 2). Note that this comparison however just approached the Bonferroni-adjusted alpha level of .017. There were differences neither between LSC and LV, $F(1, 35) = 2.08$, $p = .103$, $\eta_p^2 = .074$, nor between LSC and LTC, $F(1, 35) = 0.75$, $p = .392$, $\eta_p^2 = .021$.

For the comparisons of learned overall versus novel faces, there were main effects of familiarity for N250, $F(1, 30) = 6.54$, $p = .015$, $\eta_p^2 = .157$; LPCa, $F(1, 35) = 66.45$, $p < .001$, $\eta_p^2 = .655$; and LPCb, $F(1, 35) = 55.95$, $p < .001$, $\eta_p^2 = .615$, due to larger amplitudes for learned overall compared to novel faces. Note that for LPCa, the main effect of familiarity interacted further with laterality, $F(2, 70) = 4.11$, $p = .036$, $\eta_p^2 = .105$, $\epsilon_{HF} = .687$; Pairwise t tests for each level of laterality (i.e., C3, Cz, and C4) all revealed larger amplitudes for learned overall compared to novel faces, $t(35) > 5.40$, $ps < .001$.

In summary, veridical faces that had been learned as texture caricatures elicited numerically larger LPCb at test than did faces that had been learned as veridicals, which may highlight the importance of distinctive facial texture information during recognition. Moreover, larger N250, LPCa, and LPCb amplitudes were found for learned faces overall compared to novel faces.

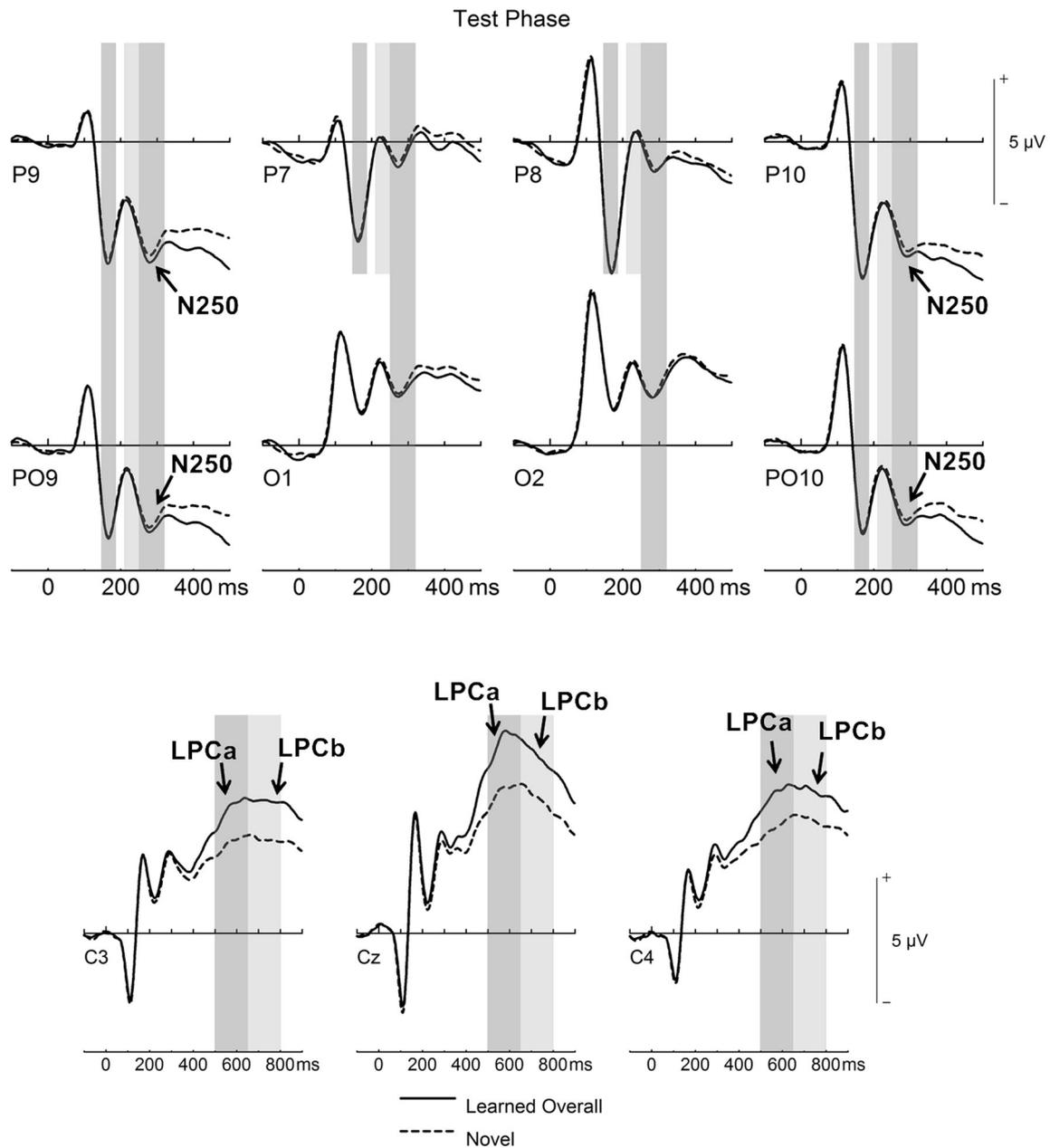


Fig. 6 Comparisons between learned overall versus novel in the test phase, exemplarily illustrated at PO10 and C4. Gray-shaded areas depict time-windows of interest: N170 (145–185 ms), P200 (210–250 ms),

N250 (250–320 ms), LPCa (500–650 ms), and LPCb (650–800 ms). Note the different time scales for N250 at occipitotemporal (top) and LPC at central (bottom) electrode sites

Discussion

To the best of our knowledge, this is the first report of caricature generalization benefits for faces selectively caricatured in either shape or texture and their underlying neural correlates. Our observation of larger benefits for veridical test faces that had been learned as either shape or texture caricatures, compared to veridical test faces that had been learned as veridicals, confirms our expectation of caricature generalization benefits. That is, learning faces with distinctive idiosyncratic shape or texture generalizes to and, importantly, *improves* recognition

of their veridical versions. Our results thus shed further light on how mental representations of faces are formed and the relative roles of distinctive facial shape and texture during that process.

Our finding of higher accuracies for veridical faces that had been learned with enhanced idiosyncratic shape or texture information compared to when they had been learned as veridicals broadly extends previous findings (Deffenbacher et al., 2000; Rodriguez et al., 2009; Stevenage, 1995). However, note that in contrast to the present experiment those studies were limited to the usage of shape caricaturing and

nonphotorealistic stimuli. Moreover, note that because we used different image viewpoints in learning and test phases for all conditions, successful face recognition in the present study must have been based on relatively robust, image-invariant representations.

ERP results in the learning phase elucidate further underlying neural processes involved during face learning. Compared to veridicals, N250 amplitudes were significantly larger for shape caricatures. Similarly, texture caricatures also elicited numerically larger N250 compared to veridicals, although this comparison was reduced to a trend after Bonferroni correction. The current finding of reliably larger N250 at learning for shape caricatures compared to veridicals potentially complements previous reports on the importance of facial shape for the initial encoding of faces, with reliable effects of texture emerging slightly later (Caharel, Jiang, Blanz, & Rossion, 2009; Itz et al., 2014; Kaufmann & Schweinberger, 2008). Moreover, largest LPCb for shape caricatures at learning also extends previous findings (Itz et al., 2014; Kaufmann et al., 2013; Kaufmann & Schweinberger, 2012; Schulz, Kaufmann, Kurt, et al., 2012; Schulz, Kaufmann, Walther, et al., 2012) and may reflect more intense processing of facial identity during learning for faces whose idiosyncratic shape has been enhanced.

Somewhat unexpectedly, at learning we found no significant effects of caricaturing for the N170. Compared to P200 and N250, this adds to the observation of more spurious N170 shape caricature effects, in particular for relatively moderate caricature levels (e.g., Itz et al., 2014; Kaufmann & Schweinberger, 2012; Schulz, Kaufmann, Kurt, et al., 2012). With respect to P200 at learning, the finding of smallest amplitudes for shape caricatures was merely a trend. We suggest that this might be due to the rating tasks at learning, which may have hampered processing of distinctive shape during the relatively early time-window of P200 (for a similar argument, see Wiese, Altmann, & Schweinberger, 2014). Moreover, in contrast to the ERP waveform morphology pattern in the test phase (see Fig. 6) and the learning phases of other studies in which there was no explicit task at learning (e.g. Itz et al., 2014; Schulz, Kaufmann, Walther, et al., 2012), the peaks between P200 and N250 in the learning phase in our data are less distinguishable, resulting from a less pronounced N250. Within the same study, prominent task effects on these components in face processing have also been demonstrated (Trenner, Schweinberger, Jentzsch, & Sommer, 2004). Thus, although further research will be needed to refine this aspect, the present differences in ERP waveform morphology between learning and test may well be because we used a nonidentity-related task at learning.

At *test*, we found larger N250 and LPC amplitudes for learned faces overall compared to novel faces, in line with our expectations and previous research (e.g., Kaufmann & Schweinberger, 2009; Tanaka et al., 2006). With respect to

learned-as condition at test, early ERP components N170 and P200 were not affected. As these early components are associated with *perceptual* stages of processing (Schweinberger, 2011), this was anticipated. Note that all test faces were veridicals and therefore did not differ systematically in terms of image properties. However, because N250 has been described as a neural marker for the strength of face familiarity (Herzmann, Schweinberger, Sommer, & Jentzsch, 2004; Kaufmann & Schweinberger, 2009; Tanaka et al., 2006), one should have expected larger N250 familiarity effects for veridical test face identities that had been learned as caricatures. Such an observation would have been in line with the present accuracy data. In this context it is important to keep in mind that N250 effects are modulated by the degree of image similarity, with familiarity and repetition effects being larger for very similar, and largest for identical images (Bindemann et al., 2008; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002). In this study we used different images for learning and testing in all conditions, but note that, due to caricaturing, in the learned-as caricature conditions there was an obvious additional source for image dissimilarity between learning and test images. This might have levelled out potentially larger N250 familiarity effects for faces learned as caricatures.

Interestingly, we found indirect ERP evidence at test for better mental representations of faces learned as caricatures, in terms of a significant main effect of learned-as condition for the later LPC time interval (650–800 ms). Pairwise comparisons suggested that this effect was mainly due to larger LPCb for learned as texture caricature (LTC) compared to learned as veridical (LV) conditions, although with a *p* value of .019 this comparison just failed to reach the critical Bonferroni-adjusted significance level of .017 and should therefore not be overinterpreted. Nevertheless, the observation is in line with other findings on the importance of texture for familiar face recognition (e.g., Burton et al., 2015; Itz et al., 2014; Russell & Sinha, 2007). As LPC has been found to reflect semantic identity processing (e.g., Schweinberger et al., 1995), at present it cannot be ruled out that our findings might reflect differences in attempted retrieval of semantic information for faces learned as texture caricatures compared to faces learned and tested as veridicals. Further research will be needed to refine this aspect.

Finally, these results carry with them important implications not only for basic face learning research but also for applied areas, for instance passport controlling (see e.g., White, Kemp, Jenkins, Matheson, & Burton, 2014). A recent study by McIntyre, Hancock, Kittler, and Langton (2013) could show that using mild levels of shape caricaturing facilitates unfamiliar face matching. Furthermore, we think that caricaturing may present a promising tool for training face recognition in individuals with poor face recognition skills (see, e.g., Irons et al., 2014; Kaufmann et al., 2013).

In conclusion, we report here for the first time caricature generalization benefits for photorealistic faces caricatured selectively in either shape or texture. Importantly, learning faces as caricatures did not merely generalize to but actually *improved* subsequent recognition of veridical counterparts. Furthermore, our ERP findings highlight the importance of facial shape for initial encoding of unfamiliar faces as well as the importance of texture for the recognition of learned faces.

Acknowledgments We gratefully acknowledge our technical assistant, Bettina Kamchen, as well as our students, Lidiia Romanova, Helge Schlüter, and Nadine Wanke, for their help with setting up the study and data collection. The authors' work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) KA 2997/3-2 to J.M.K. and S.R.S., and FOR1097 to S.R.S.

Compliance with ethical standards

Conflict of interest None.

References

- Abdi, H. (2007). The Bonferroni and Šidák corrections for multiple comparisons. In N. Salkind (Ed.), *Encyclopedia of measurement and statistics*. Thousand Oaks, CA: Sage.
- Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, 57(3), 217–239. doi:10.1016/0010-0277(95)00669-x
- Benson, P. J., & Perrett, D. I. (1991). Synthesizing continuous-tone caricatures. *Image and Vision Computing*, 9(2), 123–129. doi:10.1016/0262-8856(91)90022-h
- Berg, P., & Scherg, M. (1994). A multiple source approach to the correction of eye artifacts. *Electroencephalography and Clinical Neurophysiology*, 90(3), 229–241. doi:10.1016/0013-4694(94)90094-9
- Bindemann, M., Burton, A. M., Leuthold, H., & Schweinberger, S. R. (2008). Brain potential correlates of face recognition: Geometric distortions and the N250r brain response to stimulus repetitions. *Psychophysiology*, 45(4), 535–544. doi:10.1111/j.1469-8986.2008.00663.x
- Brennan, S. E. (1985). Caricature generator, the dynamic exaggeration of faces by computer + illustrated works. *Leonardo*, 18(3), 170–178. doi:10.2307/1578048
- Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, 51(3), 256–284. doi:10.1016/j.cogpsych.2005.06.003
- Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science*, 10(4), 482–496. doi:10.1177/1745691615583129
- Burton, A. M., & Vokey, J. R. (1998). The face-space typicality paradox: Understanding the face-space metaphor. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 51(3), 475–483.
- Caharel, S., Jiang, F., Blanz, V., & Rossion, B. (2009). Recognizing an individual face: 3D shape contributes earlier than 2D surface reflectance information. *NeuroImage*, 47(4), 1809–1818. doi:10.1016/j.neuroimage.2009.05.065
- Deffebacher, K. A., Johanson, J., Vetter, T., & O'Toole, A. J. (2000). The face typicality-recognizability relationship: Encoding or retrieval locus? *Memory & Cognition*, 28(7), 1173–1182. doi:10.3758/bf03211818
- Eimer, M. (2000a). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, 111(4), 694–705. doi:10.1016/s1388-2457(99)00285-0
- Eimer, M. (2000b). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, 11(10), 2319–2324. doi:10.1097/00001756-200007140-00050
- Hancock, P. J. B., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory & Cognition*, 24(1), 26–40. doi:10.3758/bf03197270
- Harris, A., & Nakayama, K. (2007). Rapid face-selective adaptation of an early extrastriate component in MEG. *Cerebral Cortex*, 17(1), 63–70. doi:10.1093/cercor/bhj124
- Herzmann, G., Schweinberger, S. R., Sommer, W., & Jentsch, I. (2004). What's special about personally familiar faces? A multimodal approach. *Psychophysiology*, 41(5), 688–701. doi:10.1111/j.1469-8986.2004.00196.x
- Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, 31(10), 1221–1240. doi:10.1068/p3252
- Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics*, 1, 69–82.
- Irons, J., McKone, E., Dumbleton, R., Barnes, N., He, X. M., 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). doi:10.1167/14.2.12
- Itz, M. L., Schweinberger, S. R., Schulz, C., & Kaufmann, J. M. (2014). Neural correlates of facilitations in face learning by selective caricaturing of facial shape or reflectance. *NeuroImage*, 102, 736–747. doi:10.1016/j.neuroimage.2014.08.042
- Kaufmann, J. M., Schulz, C., & Schweinberger, S. R. (2013). High and low performers differ in the use of shape information for face recognition. *Neuropsychologia*, 51(7), 1310–1319. doi:10.1016/j.neuropsychologia.2013.03.015
- Kaufmann, J. M., & Schweinberger, S. R. (2008). Distortions in the brain? ERP effects of caricaturing familiar and unfamiliar faces. *Brain Research*, 1228, 177–188. doi:10.1016/j.brainres.2008.06.092
- Kaufmann, J. M., & Schweinberger, S. R. (2009). ERP correlates of improved learning for spatially caricatured faces. *Psychophysiology*, 46, S132–S132.
- Kaufmann, J. M., & Schweinberger, S. R. (2012). The faces you remember: Caricaturing shape facilitates brain processes reflecting the acquisition of new face representations. *Biological Psychology*, 89(1), 21–33. doi:10.1016/j.biopsycho.2011.08.011
- Kaufmann, J. M., Schweinberger, S. R., & Burton, A. M. (2009). N250 ERP correlates of the acquisition of face representations across different images. *Journal of Cognitive Neuroscience*, 21(4), 625–641. doi:10.1162/jocn.2009.21080
- Latinus, M., & Taylor, M. J. (2006). Face processing stages: Impact of difficulty and the separation of effects. *Brain Research*, 1123, 179–187. doi:10.1016/j.brainres.2006.09.031
- Maurer, U., Rossion, B., & McCandliss, B. D. (2008). Category specificity in early perception: Face and word N170 responses differ in both lateralization and habituation properties. *Frontiers in Human Neuroscience*, 2. doi:10.3389/neuro.09.018.2008

- McIntyre, A. H., Hancock, P. J. B., Kittler, J., & Langton, S. R. H. (2013). Improving discrimination and face matching with caricature. *Applied Cognitive Psychology, 27*(6), 725–734.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. doi:10.1016/0028-3932(71)90067-4
- Perkins, D. (1975). A definition of caricature, and caricature and recognition. *Studies in the Anthropology of Visual Communication, 2*(1), 1–24.
- Rhodes, G. (1996). *Superportraits: Caricatures and recognition*. Hove, UK: The Psychology Press.
- Rodriguez, J., Bortfeld, H., Rudomin, I., Hernandez, B., & Gutierrez-Osuna, R. (2009). The reverse-caricature effect revisited: Familiarization with frontal facial caricatures improves veridical face recognition. *Applied Cognitive Psychology, 23*(5), 733–742. doi:10.1002/acp.1539
- Rodriguez, J., & Gutierrez-Osuna, R. (2011). Reverse caricatures effects on three-dimensional facial reconstructions. *Image and Vision Computing, 29*(5), 329–334. doi:10.1016/j.imavis.2011.01.002
- Russell, R., & Sinha, P. (2007). Real-world face recognition: The importance of surface reflectance properties. *Perception, 36*(9), 1368–1374. doi:10.1068/p5779
- Schulz, C., Kaufmann, J. M., Kurt, A., & Schweinberger, S. R. (2012). Faces forming traces: Neurophysiological correlates of learning naturally distinctive and caricatured faces. *NeuroImage, 63*(1), 491–500. doi:10.1016/j.neuroimage.2012.06.080
- Schulz, C., Kaufmann, J. M., Walther, L., & Schweinberger, S. R. (2012). Effects of anticaricaturing vs. caricaturing and their neural correlates elucidate a role of shape for face learning. *Neuropsychologia, 50*(10), 2426–2434. doi:10.1016/j.neuropsychologia.2012.06.013
- Schweinberger, S. R. (2011). Neurophysiological correlates of face perception. In A. J. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford handbook of face perception* (pp. 345–366). Oxford, UK: Oxford University Press.
- Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex, 80*, 141–153. doi:10.1016/j.cortex.2015.11.001
- Schweinberger, S. R., Pfitze, E. M., & Sommer, W. (1995). Repetition priming and associative priming of face recognition—Evidence from event-related potentials. *Journal of Experimental Psychology-Learning Memory and Cognition, 21*(3), 722–736. doi:10.1037//0278-7393.21.3.722
- Schweinberger, S. R., Pickering, E. C., Jentsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*(3), 398–409. doi:10.1016/s0926-6410(02)00142-8
- Sommer, W., Heinz, A., Leuthold, H., Matt, J., & Schweinberger, S. R. (1995). Metamemory, distinctiveness, and event-related potentials in recognition memory for faces. *Memory & Cognition, 23*(1), 1–11. doi:10.3758/bf03210552
- Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: An event-related potential study. *Neuroreport, 19*(5), 583–587.
- Stevenage, S. V. (1995). Can caricatures really produce distinctiveness effects? *British Journal of Psychology, 86*, 127–146.
- Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: The N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience, 18*(9), 1488–1497. doi:10.1162/jocn.2006.18.9.1488
- Trenner, M. U., Schweinberger, S. R., Jentsch, I., & Sommer, W. (2004). Face repetition effects in direct and indirect tasks: An event-related brain potentials study. *Cognitive Brain Research, 21*(3), 388–400.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 43*(2), 161–204.
- Valentine, T., & Bruce, V. (1986). Recognizing familiar faces - the role of distinctiveness and familiarity. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 40*(3), 300–305. doi:10.1037/h0080101
- Valentine, T., Lewis, M. B., & Hills, P. J. (2014). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*. doi:10.1080/17470218.2014.990392
- Vokey, J. R., & Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition, 20*(3), 291–302. doi:10.3758/bf03199666
- White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLOS ONE, 9*(8). doi:10.1371/journal.pone.0103510
- Wiese, H., Altmann, C. S., & Schweinberger, S. R. (2014). Effects of attractiveness on face memory separated from distinctiveness: Evidence from event-related brain potentials. *Neuropsychologia, 56*, 26–36. doi:10.1016/j.neuropsychologia.2013.12.023
- Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (2014). The neural signature of the own-race bias: Evidence from event-related potentials. *Cerebral Cortex, 24*(3), 826–835. doi:10.1093/cercor/bhs369