Nonlinear transduction of emotional facial expression

Article

Accepted Version

Gray, K. L.H., Flack, T. R., Yu, M., Lygo, F. A. and Baker, D. H. (2020) Nonlinear transduction of emotional facial expression. Vision Research, 170. pp. 1-11. ISSN 0042-6989 doi: https://doi.org/10.1016/j.visres.2020.03.004 Available at http://centaur.reading.ac.uk/89417/

To link to this article DOI: http://dx.doi.org/10.1016/j.visres.2020.03.004

Publisher: Elsevier

Nonlinear transduction of emotional facial expression

Katie L.H. Gray[1], Tessa R. Flack[2], Miaomiao Yu[3], Freya A. Lygo[3] & Daniel H. Baker[3,4]

1. School of Psychology and Clinical Language Sciences, University of Reading, Reading, RG6 6BZ, UK

2. School of Psychology, University of Lincoln, Brayford Pool, Lincoln, LN6 7TS, UK

3. Department of Psychology, University of York, Heslington, York, YO10 5DD, UK

4. York Biomedical Research Institute, University of York, Heslington, York, YO10 5DD, UK

Abstract

To create neural representations of external stimuli, the brain performs a number of processing steps that transform its inputs. For fundamental attributes, such as stimulus contrast, this involves one or more nonlinearities that are believed to optimise the neural code to represent features of the natural environment. Here we ask if the same is also true of more complex stimulus dimensions, such as emotional facial expression. We report the results of three experiments combining morphed facial stimuli with electrophysiological and psychophysical methods to measure the function mapping emotional expression intensity to internal response. The results converge on a nonlinearity that accelerates over weak expressions, and then becomes compressive for stronger expressions, similar to the situation for lower level stimulus properties. We further demonstrate that the nonlinearity is not attributable to the morphing procedure used in stimulus generation. A preprint of this work is available at: https://doi.org/10.31234/osf.io/svw8q

Keywords: emotional expressions; nonlinear transduction; SSVEP; psychophysics; morphing.

1. Introduction

Facial expressions are communicative tools; they signal an individual's emotional state and motivation, and provide us with a wealth of information in social contexts (Adolphs, 2002; Öhman, 2002). An expression can range from very subtle to very intense, and previous work has used morphing software to parametrically manipulate emotional intensity within faces of the same identity (Blair, Colledge, Murray, & Mitchell, 2001; Harris, Young, & Andrews, 2012; Hess, Blairy, & Kleck, 1997). But how do changes in stimulus intensity map onto changes in the brain's response to, and our perception of, another's face? Despite the importance of this question for our understanding of perceived emotion, the precise mapping is currently unclear.

Nonlinearities in the neural representation of low-level image features are very well established. The brain responds to image contrast (defined as the luminance difference between the brightest and darkest parts of an image, scaled by the mean luminance) according to a saturating nonlinearity, that accelerates at intermediate contrasts, and becomes shallow at higher contrasts. This pattern is consistent across measurements using psychophysical contrast discrimination, matching and scaling paradigms (Kingdom, 2016; Legge & Foley, 1980), functional magnetic resonance imaging (fMRI; Boynton, Demb, Glover, & Heeger, 1999), electroencephalography (EEG; Campbell & Kulikowski, 1972; Tsai, Wade, & Norcia, 2012), single- and multi-unit recording (Albrecht & Hamilton, 1982; Busse, Wade, & Carandini, 2009; Ohzawa, Sclar, & Freeman, 1982) and optical imaging using voltage sensitive dyes (Reynaud, Barthélemy, Masson, & Chavane, 2007).

Measuring neural responses to higher order stimulus properties (such as facial expression) is possible using a fast periodic visual stimulation (FPVS) technique, which induces oscillations in the EEG signal at specific frequencies. In this paradigm, 'oddball' target stimuli (e.g. faces bearing an expression, or of a specific identity) are interleaved within a sequence of base stimuli (e.g. neutral faces, or faces of a different identity) at a specific temporal frequency. If the target can be discriminated, responses are evident at harmonics of the oddball frequency (Braddick, Wattam-Bell, & Atkinson, 1986; Liu-Shuang, Norcia, & Rossion, 2014). Most previous studies have used high intensity expressions and made comparisons across different configurations (e.g. upright and inverted; Coll, Murphy, Catmur, Bird, & Brewer, 2019; Dzhelyova, Jacques, & Rossion, 2017). However, by parametrically varying the intensity of emotional expression in the oddball stimulus, an 'emotion-response function' (analogous to a contrast-response function) can be measured. This directly reveals the transfer function between facial expression intensity and neural response. One recent study (Leleu et al., 2018) has reported such an experiment, and shown evidence of nonlinear components in the emotion-response function.

The perceptual consequences of neural nonlinearities can also be measured in a variety of ways. For stimulus levels around detection threshold, the slope of the psychometric function (the function relating stimulus intensity to accuracy in a two-alternative-forced-choice detection task) depends on the underlying transducer nonlinearity in that region of stimulus space (assuming no uncertainty about the task). A linear system will result in a shallow psychometric function (Weibull β values around 1.3, see Meese & Summers, 2012; Pelli, 1985; Tyler & Chen, 2000), whereas accelerating nonlinearities produce steeper slopes. There is some evidence from recent work (Marneweck, Loftus, & Hammond, 2013) of slopes with β >

1.3 for discriminating four distinct emotional expressions from neutral, though deviation from linearity was not formally assessed.

A complementary approach to characterize signal processing is to use a discrimination paradigm, in which a participant's ability to detect differences in magnitude is measured at a range of starting ('pedestal') levels (Nachmias & Sansbury, 1974). Relative to detection in the absence of a pedestal, weak pedestal levels can reduce the target level required to reach threshold performance (facilitation), whereas strong pedestal levels can increase thresholds (masking). The combination of these effects creates a characteristic 'dipper' shaped function (Legge & Foley, 1980) when threshold is plotted against pedestal level, that is determined by the gradient (steepness) of the underlying nonlinearity. A linear system would not produce either the facilitation or masking effects, and thresholds should remain constant regardless of pedestal level. Dipper functions have been reported for a range of sensory cues, including motion (Gori, Mazzilli, Sandini, & Burr, 2011), blur (Watt & Morgan, 1983), depth (Georgeson, Yates, & Schofield, 2008), texture (Morgan, Chubb, & Solomon, 2008), duration (Burr, Silva, Cicchini, Banks, & Morrone, 2009), loudness (Raab, Osman, & Rich, 1963), and amplitude modulation (Nelson & Carney, 2006), suggesting that the underlying nonlinearity is a common property of perceptual systems.

One previous study has applied a similar paradigm to investigate the representation of facial identity. Dakin and Omigie (2009) measured identity-strength discriminability of faces using an odd-one-out paradigm. They morphed between an average identity face and a full identity face in a number of steps. They then presented three faces: two identical faces (containing the pedestal level of identity), and one face containing the pedestal identity with an additional increment of identity. They repeated this at a number of different identity pedestal-levels, measuring sensitivity at each level. When plotting threshold against pedestal identity, they found evidence for shallow dipper-shaped functions, suggestive of a nonlinearity in the representation of identity. However, these functions typically lacked the masking region found for contrast (the dipper 'handle'). Work by Marenweck, Loftus and Hammond (2013) reports discrimination for emotional expressions, but the pedestal level was not fixed within a condition, making interpretation difficult. A primary aim of the present study is to investigate whether emotional expression intensity is also subject to a process of nonlinear transduction by measuring thresholds for expression discrimination at a range of pedestal levels.

Here we report the results of three experiments. In the first we use an EEG paradigm to measure neural responses to facial expressions in order to map out an emotion-response function. In the second we measure the slope of the psychometric function for an expression detection task. Finally, we assess the discriminability of emotional expressions from a range of baseline (pedestal) levels. The results give a comprehensive picture of how expression intensity information is processed to form an internal representation of others' emotional states. We find evidence of a nonlinear transduction process similar to that reported for other variables, which accelerates at low expression levels, and becomes shallower for more intense expressions.

2. Methods

2.1 Participants

Twenty-four adult participants completed the EEG and detection experiments (M_{age} = 23; SD = 5.29; 5 males), and six participants completed the discrimination experiment (1 male). All had normal or corrected-to-normal visual acuity. All experiments were approved by the ethics committee of the Department of Psychology at the University of York, and written informed consent was obtained from all participants.

2.2 Apparatus and stimuli

All stimuli were derived from greyscale male and female faces taken from the NimStim face set (Tottenham et al., 2009), depicting 6 basic emotional expressions (angry, fear, happy, sad, surprise, and disgust; Ekman & Friesen, 1971). In the EEG and detection experiments, we used 16 female and 22 male identities, having a variety of racial backgrounds. For each identity, we used a program (developed by Adams, Gray, Garner, & Graf, 2010) to morph between neutral and an emotional expression in 6 steps, creating 7-levels of emotional intensity: 0, 6, 12, 24, 48, 96 and 144% (e.g. Calder et al., 2000; Calder, Young, Rowland, & Perrett, 1997). For the discrimination experiment, we also created an averaged identity for each gender (based on 19 female and 23 male exemplars), and then morphed between neutral and 150% expression in 0.5% steps. External features (i.e. hair and ears) were removed from all faces using an elliptical mask blurred by a cosine function. All stimuli were equated for mean luminance and root-mean-square contrast.

In the EEG experiment, brain activity was recorded from 64 scalp locations laid out according to the 10/20 system in a WaveGuard cap (ANT Neuro, Netherlands). We also monitored blinks through bipolar electro-oculogram electrodes placed above and below the left eye. Signals were amplified and digitised at 1kHz and recorded using the ANT Neuroscan software (ANT Neuro, Netherlands). Stimuli were presented using a gamma corrected VIEWPixx display (VPixx Technologies Inc., Quebec, Canada) with a resolution of 1920x1200 pixels, a mean luminance of 50cd/m^2, and a refresh rate of 120Hz, controlled by an Apple Macintosh computer. Trigger codes were sent from the VIEWPixx device to the EEG amplifier using a 25-pin parallel port to identify each condition and record stimulus onset times. The PsychToolbox routines (Brainard, 1997) running in MATLAB were used to control the display hardware and send triggers. The same display hardware was used in the detection experiment, but EEG activity was not recorded. In the discrimination experiment, stimuli were centrally presented on a gamma corrected 21-inch Iiyama VisionMaster Pro 510 monitor with a mean luminance of 32cd/m^2 and a resolution of 1152x768 pixels, driven at 75Hz by an Apple Macintosh computer.

2.3 Procedures

EEG experiment: Sequences of faces were presented for trials of 60 seconds duration. Faces subtended approximately 8x12 degrees of visual angle at the viewing distance of 57cm, and were presented against a grey background with a central black fixation cross. The contrast of the faces was modulated between 0 and 100% according to a 5Hz sine wave (see Figure 1a). The identity of the face was changed at the minimum of each period (when the contrast was zero), resulting in a seamless stream of different identities. In this paradigm, each face stimulus was presented for 200ms, but because contrast was 0 at the face onset and offset, each face was visible for around 180ms. All stimuli had a neutral expression, except for an ‘oddball’ stimulus presented every fifth cycle (i.e. at 1Hz; see Figure 1a). This stimulus had a randomly selected expression on each presentation, at a specific morph level that was constant throughout the trial. Similar timings have been used previously with face stimuli (Liu-Shuang et al., 2014; Rossion, Prieto, Boremanse, Kuefner, & Van Belle, 2012) and appear to be a good compromise between potential floor and ceiling effects (i.e. too fast to allow isolation of each individual response, or too slow to give large face-selective responses). Participants were asked to fixate on a central cross for the duration of the trial and try to minimise blinking; there was no behavioural task. Each block consisted of eight trials; one for each morph level, plus an inversion condition using the 96% expression, but with all faces rotated through 180 degrees. There was an inter-trial interval of 8 seconds. Each participant completed four repetitions, taking around 40 minutes in total.

Detection experiment: We used a two-interval forced choice procedure that was designed to closely mirror the temporal properties of the EEG experiment. Participants were presented with two sequential streams of faces; a target stream containing a single emotional face embedded within 8 neutral distractors, and a null stream containing only neutral faces. The target face always appeared on the fifth cycle (the midpoint of the target stream; see Figure 1b). The target and distractors were random identities, and the same identity was never repeated on two adjacent cycles. The two streams were separated by 500ms. Participants were asked to detect which stream contained the emotional target, and indicated their responses using a mouse. Target intensity, target expression, and target interval were randomised across trials. There were 480 trials (60 per emotional intensity condition, including 60 trials for the inversion condition at the 24% morph level), separated into 5 blocks, taking around 40 minutes to complete.

Discrimination experiment: We used a two-interval forced choice procedure; on each trial, a face (subtending 10x16 degrees at the viewing distance of 57cm) was presented centrally for 100ms in each of two intervals, separated by 400ms. One face had its expression set at the pedestal level (the null stimulus; pedestal levels were 0, 15, 30, 45, 60 and 75%), the other face had its expression set at the pedestal level plus an increment (the target stimulus). Participants indicated which interval contained the face with the strongest expression intensity (i.e. the target) using a mouse. In additional conditions, pedestal and target stimuli were applied to different halves of the face; the results of these conditions will be reported in a subsequent publication. Stimuli were surrounded by a black square, and divided horizontally by a black line. The purpose of the black line was to mask luminance discontinuities caused by combining upper and lower face halves from different expression intensities in some conditions, and is consistent with standard composite effect procedures (Rossion, 2013). The gender of the face was chosen randomly on each trial (with equal probability), but was the same across the null and target intervals. The expression was constant across the null and target intervals, but was chosen at random on each trial in the main experiment. On each trial, the level of the target increment was selected using a staircase procedure (three-down, one-up, step size of 2.5%) that terminated after the lesser of 70 trials or 12 reversals. Participants received auditory feedback on the accuracy of each response. The main experiment took around 4.5 hours to complete for each participant, and consisted of around 8000-9000 trials per participant (of which around ¼ are reported here). We also ran a control experiment for a restricted set of pedestal levels, in which the expression was fixed within a block.

2.4 Data Analysis

EEG experiment: We took the Fourier transform of the EEG waveform (i.e. transformed the responses from the time domain to the frequency domain) from each electrode for the 60 seconds during which stimuli were presented. There was a strong response from occipital electrodes at the baseline frequency (5Hz) in all conditions, reflective of the general change in contrast (and other image properties, such as identity) of the stimuli at this rate. Our measure of interest was the amplitude at harmonics of the oddball frequency (1Hz), as this measure is specific to emotional expression. To calculate the responses to the oddball stimuli, we took the coherent average across repetitions and participants at 2, 3 and 4Hz, and then averaged the amplitudes across these three frequencies to provide a single measure. We did not include responses at 1Hz, as these were not distinguishable from the high noise levels in this region of the spectrum (see Figure 1c), consistent with previous studies (Liu-Shuang et al., 2014). We also excluded responses at and above the baseline frequency (>=5Hz), as these are difficult to interpret given the strong contribution from the baseline flicker component.

Detection and discrimination experiments: Individual thresholds were estimated from each participant's responses (as well as the pooled data in the detection experiment) by fitting a cumulative Weibull function using the *quickypsy* package in *R* (Linares & López-Moliner, 2016). We defined threshold as the morph intensity required to reach 81.6% correct (i.e. the balance point of the Weibull function), and the slope as the β parameter of the fit.

Data and code availability: Primary analyses were performed in *R*. Analysis scripts and raw data are available at: http://dx.doi.org/10.17605/OSF.IO/8MS4Y

3. Results

3.1 The emotion-response function is nonlinear

In our first experiment, we measured the neural response to stimuli of different emotional intensities using a steady-state FPVS EEG paradigm, in a group of 24 adults. Streams of face images with random identities were presented at 5Hz, with every fifth 'oddball' image bearing a randomly chosen emotion, and the remainder being neutral (see Figure 1a). When the oddball faces were also neutral (i.e. had a 0% expression morph level) there were clear responses only at the carrier modulation frequency of 5Hz (see Figure 1b). When the oddball faces carried a strong expression, responses were also evident at harmonics of the oddball frequency (i.e. multiples of 1Hz, see Figure 1c), and were strongest over parieto-occipital electrodes in the right hemisphere. These responses increased monotonically with morph level at each of the first three harmonics (2, 3 and 4Hz), as shown by the lines in Figure 1d, and their average (orange-filled circles in Figure 1d). Consistent with previous work (Dzhelyova et al., 2017), inverting all images in the stream generated a much weaker expression-specific response, as shown by the green symbol in Figure 1d (paired t-test; t=5.29, df=23, p=0.000023, d=1.1, BF=1025).

To assess the linearity of these data, we replotted the average across the first three harmonics on log-log axes (see Figure 1e). The best fit regression line to these data had a slope of 0.73, and the upper bound of a bootstrapped 95% confidence interval on this slope estimate was also below 1 (lower CI = 0.54; upper CI = 0.91). This is evidence of a compressive nonlinearity, equivalent to $y = x^{0.73}$, where x is morph level.

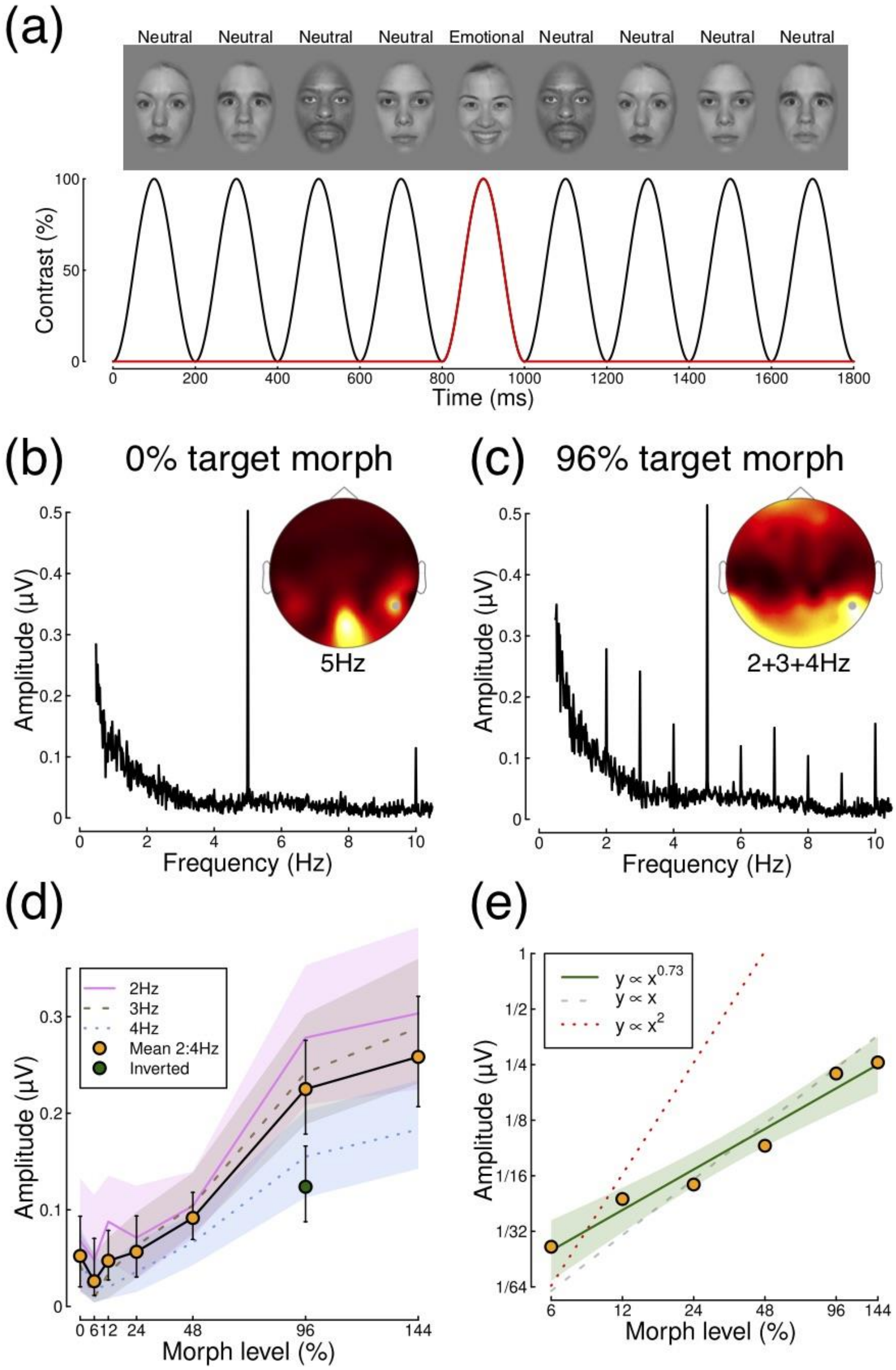

Figure 1: Neural SSVEP responses are lateralised and nonlinear. Panel (a) represents the stimuli presented during a brief (1.8s) period of an extended (60s) trial. Stimulus contrast was sinusoidally modulated at 5Hz, with the

face image changed every 200ms at the trough of the modulation. An 'oddball' emotional face was presented every 5 cycles, at a rate of 1Hz. Panel (b) shows the Fourier spectrum in the condition where the oddball stimuli were also neutral, averaged across all participants (N=24). A strong response is evident at the modulation frequency (5Hz), which is maximal at the occipital pole, with additional activity at more lateral sites. The spectrum is derived from electrode P8, shown by the grey point. Panel (c) shows the Fourier spectrum for a 96% target morph level. Here additional peaks in the spectrum are evident at integer frequencies. Panel (d) shows emotion-response functions at individual frequencies (2, 3 and 4Hz) and their average (orange points). Shaded regions and whiskers represent bootstrapped 95% confidence intervals across participants. Panel (e) shows the average data replotted on log-log axes. Dashed and dotted lines show canonical predictions for a linear system (dashed) and a squaring nonlinearity (dotted). The solid green line shows the best fit regression line in logarithmic units, which has a slope of 0.73, with the green shaded region giving 95% confidence intervals of the regression line.

3.2 A nonlinear psychometric function for emotion detection

We next sought to measure the psychometric function for detection of emotional expressions as a function of morph level. We based the stimulus sequence on that used in the SSVEP experiment, and presented two sequences of 9 face images, each lasting 1.8 seconds (see Figure 1a). One sequence comprised only neutral faces, and the other contained an emotional face as the fifth image. Participants indicated which sequence they believed contained the emotional face. Performance increased monotonically as a function of morph level, from chance performance at low morph levels (0-12%), reaching near ceiling performance for morph levels of 96 and 144% (see Figure 2a). Again, there was an inversion effect (see green point in Figure 2a), which reduced accuracy from 0.66 to 0.59 when the faces were presented upside-down (paired t-test; t=3.19, df=23, p=0.004, d=0.65, BF=10.28).

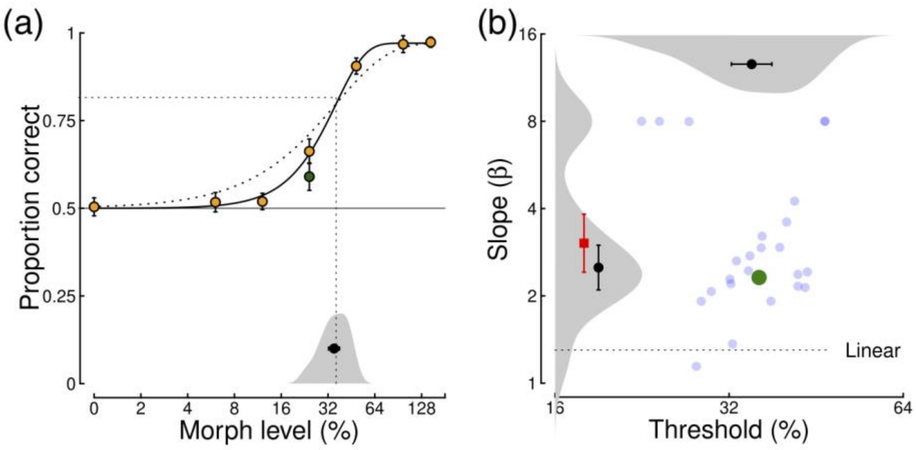

Figure 2: Nonlinear psychometric functions for detection of emotional expression. Panel (a) shows the group average psychometric function (N=24), along with the best fitting Weibull function (black solid curve). The grey shaded region at the foot shows the distribution of individual thresholds, along with the mean (black point). The black dotted curve is a Weibull function with the same threshold, but a slope of $\beta = 1.3$, showing the prediction for a linear system. Panel (b) shows individually fitted thresholds and slopes (blue points), along with the fit to the group average data (green). Grey shaded regions show distributions for each parameter, along with their means across participants (black points). For slope values, the red square is the mean with the 4 outliers at $\beta = 8$ included, and the black point shows the mean with the outliers excluded. The dotted black line at $\beta = 1.3$ gives the prediction for a linear system. Error bars in both panels show 95% confidence intervals.

We fitted a cumulative Weibull function to the group averaged psychometric function (see solid curve in Figure 2a), and also to the functions for each individual participant (N=24), to estimate the threshold and slope. The group average threshold at 81.6% correct occurred at a morph level of 31.0%. This agreed well with the mean of the individual thresholds, which was 30.9%. The psychometric slope for the group averaged data was $\beta = 2.31$, substantially above the slope expected for a linear system of $\beta = 1.3$ (assuming no uncertainty). A psychometric function with a slope of $\beta = 1.3$ is shown by the dotted curve in Figure 2a, and is a poor fit to the data. Because slope values can sometimes be underestimated for group data if individual participants have different thresholds (see e.g. Wallis, Baker, Meese, &

Georgeson, 2013), we also assessed the slope values of individual fits (see Figure 2b). The geometric mean psychometric slope across the group was β = 2.9, which was also above the linear prediction of β = 1.3 (t=7.42, df=23, p<0.001, d=1.51, BF=101258). Four fits returned a slope at the upper bound of the permitted values (β = 8). When these participants were excluded, the geometric mean slope reduced to β = 2.4, which was still significantly steeper than β = 1.3 (t=8.88, df=19, p<0.001, d=1.98, BF=396167).

The slope value of $\beta \approx 2.4$ corresponds to an effective transduction exponent of approximately 2.4/1.3 = 1.85. How can we reconcile this apparently accelerating nonlinearity around detection threshold with the compressive nonlinearity implied by our EEG data? One likely explanation is that the SSVEP paradigm was not sufficiently sensitive to detect responses in the sub-threshold range of morph levels (morph levels below 48% did not generate responses that were reliably above the noise floor, see Figure 1d). On the other hand, psychophysical performance had almost asymptoted by this morph level (see Figure 2a). The two results can therefore be considered complementary, as they reveal the nonlinearities operating in different ranges of the stimulus continuum. This is also consistent with other cues, such as contrast, which feature an accelerating nonlinearity around threshold and a compressive regime at higher stimulus intensities (e.g. Legge & Foley, 1980; Meese, Georgeson, & Baker, 2006). This combination of nonlinearities should result in a 'dipper' function for emotional expression intensity discrimination; our final experiment investigates this prediction.

3.3 A 'dipper' function for emotion discrimination

We measured emotion discrimination functions in six participants using a two-interval forced choice paradigm. To avoid the potentially complicating factors of temporal and identity uncertainty that might stem from the stimulus presentation sequences used in the previous experiments, we simplified the paradigm in two ways. First, only a single face was presented on each interval of a trial. Second, this face was an averaged identity, created by morphing either male or female faces (see Figure 3a,b for examples). We measured discrimination at a range of pedestal levels using a staircase method, and then fitted psychometric functions (see Figure 2a) to estimate thresholds. A linear system should produce a completely flat function for discrimination paradigms, where the pedestal level has no effect on threshold; any modulation of thresholds is therefore evidence of nonlinear processing.

Thresholds at six pedestal morph levels are shown in Figure 3c. For a pedestal level of 0%, the task is one of emotion detection. On average, participants required morph levels of around 29% to reliably detect (at 81.6% correct) the interval containing an emotional face (leftmost point in Figure 3c). This compared closely with thresholds in the earlier experiment (mean of 31% morph level) using the method of constant stimuli with a different stimulus set and temporal sequence. For weak pedestal expressions (15% morph level) sensitivity to the target increment improved (i.e. thresholds decreased) by around a factor of 1.6, showing evidence of facilitation from the pedestal. At higher pedestal levels a masking effect occurred, whereby increment thresholds were higher than without a pedestal. This pattern was evident for each individual participant (red lines in Figure 3c). Overall, there was a substantial effect of pedestal level on threshold ($F(5,25)=23.49$, $p<0.001$, $\eta^2=0.75$, BF=7758025) that was driven by thresholds in the 0% pedestal condition being significantly higher than in the 15% pedestal condition ($t(5)=5.68$, $p=0.002$, $d=2.32$, BF=20.72), and lower than in the 60% and 75% pedestal conditions ($t(5)=-3.33$, $p=0.021$, $d=1.36$, BF=3.98; $t(5)=-3.63$, $p=0.015$, d=1.48, BF=5.06, respectively). The slope of the rising limb of the dipper handle (estimated using linear regression over the highest four pedestal contrasts) was 0.57 (95% CIs: 0.41, 0.73).

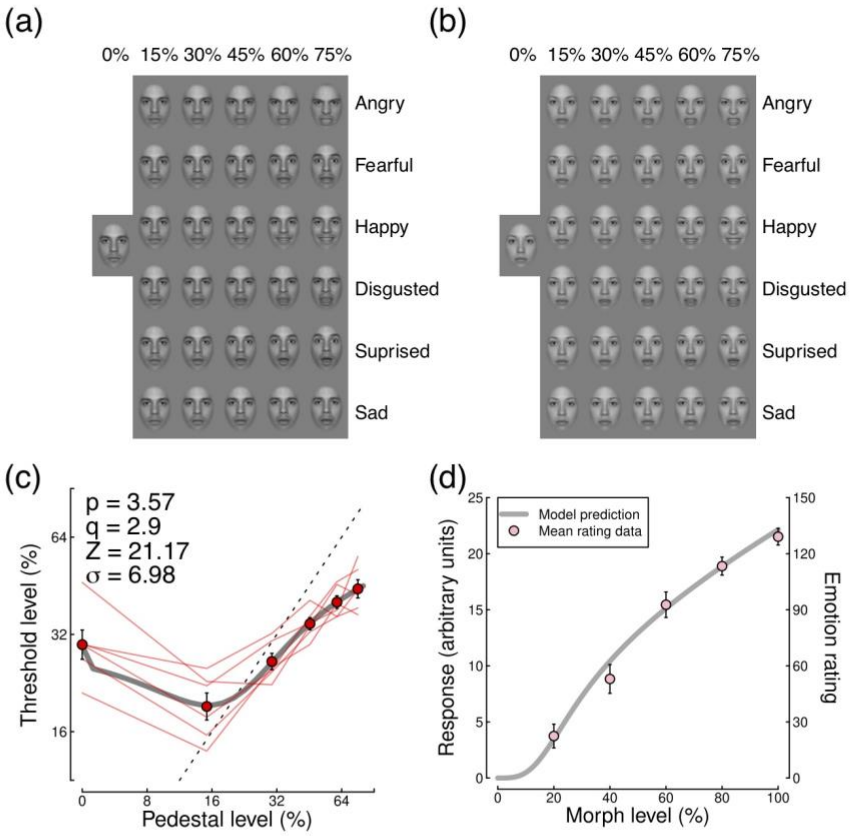

Figure 3: A dipper function for emotion discrimination. Panels (a,b) show example morphed facial stimuli for 6 expressions at the pedestal morph levels, for male (a) and female (b) averaged identities. Panel (c) shows the emotion discrimination function for individual participants (N=6, red lines) and their average (points; error bars show ±1SE). The grey curve shows the best model fit (see text for details), and the dashed oblique line has unit slope. Panel (d) shows the underlying emotion response function implied by the model fitted to the data in (c). Pink points replot the averaged data of Hess et al. (1997).

We fitted the average data with a standard nonlinear transducer function (Legge & Foley, 1980) with four free parameters. The response to a face of a given intensity level (I) is given by,

$$f(I) = \frac{I^p}{Z^q + I^q}, \qquad (1)$$

where *p*, *q*, and *Z* are free parameters. Thresholds are determined by calculating the increment level that satisfies *f*(*pedestal+increment*) = *f*(*pedestal*) + *σ*, where *σ* is a further free parameter that represents internal noise in the system. We determined best fitting parameters using a downhill simplex algorithm that minimised the least-squares error between data and model predictions. The best fitting curve is shown in Figure 3c, with parameters in the upper left corner. With four free parameters, the model provides an excellent description of the data, yielding an RMS error of 0.05dB.

In Figure 3d we plot the underlying transducer nonlinearity (the output of equation 1 for a range of inputs) using the parameters derived from the fit in Figure 3c. The function has a steep region around morph levels between 10% and 40% (i.e. around detection threshold), but becomes shallower (compressive) at higher morph levels. This function represents the way in which stimuli of different emotional intensities are mapped onto an internal response scale, and shares several common features with the rating scale data of Hess et al. (1997), most especially the shallowing at higher intensity levels. The points in Figure 3d replot the data from Hess et al. (1997) averaged across expression (anger, disgust, happiness and sadness) and face gender. It is clear that the data show extremely good correspondence with the predictions of the model, with no additional free parameters required (though note that the y-axes are scaled independently for the data points and the curve). In particular, the slope of the function at high intensity levels accurately predicts that observed in the data.

3.4 Uncertainty reduction cannot explain the facilitation effect

An alternative explanation for facilitation effects that does not require a nonlinear transducer is uncertainty reduction (Pelli, 1985). Under this account, at detection threshold an observer is uncertain about which mechanisms to monitor and performs poorly. When the pedestal is added, this helps the observer determine which mechanisms (or features of the stimulus) to attend to, and performance improves (facilitation). Because the facial expressions shown in our experiments were determined randomly on each trial, we wondered if the facilitation effects could be explained by expression uncertainty. To test this, we conducted a control experiment (on five participants) in which we blocked trials by emotion. Participants were explicitly told at the beginning of a block of trials which emotion would be presented. All other experimental parameters were the same as for the main dipper experiment.

Results for this control experiment are presented in Figure 4. For all expressions, a facilitation effect was still observed at 15% pedestal level. There were variations in sensitivity across expressions (circles; see also Marneweck et al., 2013); in particular thresholds were somewhat higher for sad expressions (pink symbols) than they were for other expressions. The average thresholds from the blocked conditions (black lines) were slightly lower than those from the interleaved method used in the main experiment (red lines). A 2 (pedestal level) x 2 (blocking condition) ANOVA showed a main effect of pedestal level ($F(1,4)=47.79$, $p=0.0023$, $\eta_p^2=0.92$) but no effect of blocking condition ($F(1,4)=3.63$, $p=0.13$) or interaction effect ($F(1,4)=1.44$, $p=0.30$). We can therefore conclude that uncertainty effects were minimal for our paradigm, and the dipper effect we report can be most straightforwardly explained by a transducer nonlinearity.

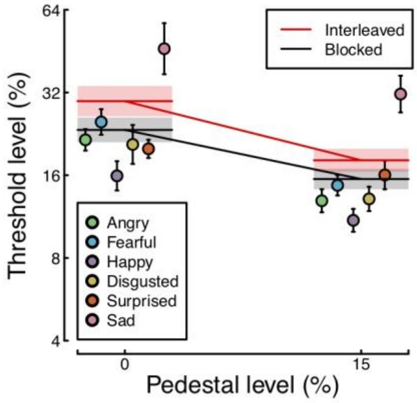

Figure 4: Facilitation effects occur for individual emotional expressions. Circles show thresholds for individual emotions for the blocked control conditions, and the black horizontal bars give their average. The red horizontal bars represent analogous conditions from the main experiment for the five participants who completed the control experiment. Error bars and shaded regions show ±1SE across participants (N=5).

4. Discussion

We have demonstrated a nonlinear mapping between the facial expression intensity in a stimulus and the internal response magnitude evoked by that stimulus. Across three experiments, we find that the nonlinearity is extremely similar to that reported for more basic visual dimensions such as contrast. Responses are negligible at low intensities, rise steeply at intermediate intensities around threshold, and exhibit a shallower, compressive portion at high intensities (Figure 3d). The nonlinearity produces facilitation and masking effects in an expression discrimination task, leading to a ‘dipper’ function similar to those reported for a range of other sensory cues, and accurately predicts rating data from a previous study.

What is the purpose of this nonlinear transduction process for expression intensity? One explanation for similar phenomena in contrast transduction (e.g. contrast gain control; Carandini & Heeger, 2012; Heeger, 1992) is that they focus the greatest sensitivity in the region of intensities most commonly experienced in the environment, or that is of most use to the organism. In everyday social interactions, individuals rarely display extremes of emotion with the intensities associated with our 100% morphs (middle image in Figure 1a). Instead, most of the expressions we encounter in real life are weaker, and perhaps quite fleeting. Yet it is crucially important that we are able to detect and discriminate changes in these expressions to gauge the emotional states of our conspecifics. Therefore a mechanism that is most sensitive to changes in weak emotions is likely to have been most useful during human evolution. It is also likely that adaptation to emotional expressions (e.g. Adams et al., 2010; Butler, Oruc, Fox, & Barton, 2008; Fox & Barton, 2007; Juricevic & Webster, 2012; Webster, Kaping, Mizokami, & Duhamel, 2004; Winston, Henson, Fine-Goulden, & Dolan, 2004) serves to maintain this sensitivity even when individuals display more extreme levels of emotion on average.

The use of stimuli that are morphed along continua of expression or identity has become increasingly common in face processing research. Yet some such studies implicitly assume that linear steps in the morph space should correspond to linear differences in perception (Blair et al., 2001; Orgeta & Phillips, 2008; Rotshtein, Henson, Treves, Driver, & Dolan, 2005). Our data, along with those of others (Dakin & Omigie, 2009; Hess et al., 1997; Leleu et al., 2018), indicate that this assumption is incorrect. Our decision to use a neutral expression as a baseline condition was arbitrary (see Young et al., 1997), and we anticipate that similar results would be obtained when morphing between two emotional expressions (see Chen, Pan, & Chen, 2014 for preliminary evidence of this), or with other facial attributes associated with character traits such as trustworthiness and dominance (Oosterhof & Todorov, 2008). This suggests that multidimensional 'face space' accounts (e.g. Russell & Bullock, 1986; Valentine, 1991) must become more complex than previously proposed, because of the need to incorporate nonlinear processes that will distort the space (Tanaka, Giles, Kremen, & Simon, 1998).

Category boundary effects for both emotional expression (Calder, Young, Perrett, Etcoff, & Rowland, 1996; Etcoff & Magee, 1992) and facial identity (Beale & Keil, 1995) have been widely reported, and can be considered a severe form of nonlinearity. Categorical processing is typically defined by a rapid transition between categories (e.g. neutral and happy expressions, or between two identities), and more similar perception or neural activity within rather than between categories, even for comparable physical changes to the stimulus (Rotshtein et al., 2005). We suspect our finding of a steep psychometric function for detection (Figure 2), and a transducer that accelerates and then compresses (Figure 3d) might meet the criteria often used for identifying categorical perception, and think it unlikely that our data could discriminate between these two explanations. However, we note that category effects are formally equivalent to high-threshold theory, which has been widely discredited for low-level cues in favour of a signal detection theory approach (Nachmias, 1981; Tyler & Chen, 2000). Characterising the underlying nonlinearity, as we have done here, offers greater explanatory and predictive power (e.g. Figure 3d) than positing a binary category boundary.

Alternatively, it may be that different brain regions contain categorical and continuous representations of emotional expression, with evidence that cortical regions in the temporal lobe contain a continuous representation, whereas subcortical structures including the amygdala contain a categorical representation (Harris et al., 2012). Since subcortical structures are too deep for EEG to probe directly, our SSVEP signals most likely originate in cortical regions from which EEG activity can be detected, explaining the continuous response we report (see Figure 1e). On the other hand, cortical responses might also relay activity from subcortical regions, though presumably further processing would be applied in cortex that might change the nature of the response.

4.1 Alternative metrics still support nonlinear processing

In all our experiments we used a morphing technique to generate intermediate levels of emotional expression. The morphing process produces a linearly increasing sequence of expressions, but it manipulates the images geometrically in two dimensions, which could introduce nonlinearities into the low level image features. In principle the apparently neural nonlinearities we measure experimentally could be inherited from the stimuli if participant responses were based on cues other than expression. We quantified this in two ways to investigate whether image nonlinearities might be responsible for the apparently nonlinear processing that we report. First, we measured the average absolute difference between pixels in each successive morphed face image (the square root of the mean squared difference produced a very similar result). This gives an aggregate measure of how local luminance changes as a function of morph level, and shows evidence of a mild nonlinearity (see Figure 5a). Second, we measured the average absolute amplitude difference at each orientation and spatial frequency in the Fourier transform of the images. This gives an indication of how the global spectral content of the images changes as a function of morph level, and shows a more profound nonlinearity (see Figure 5e).

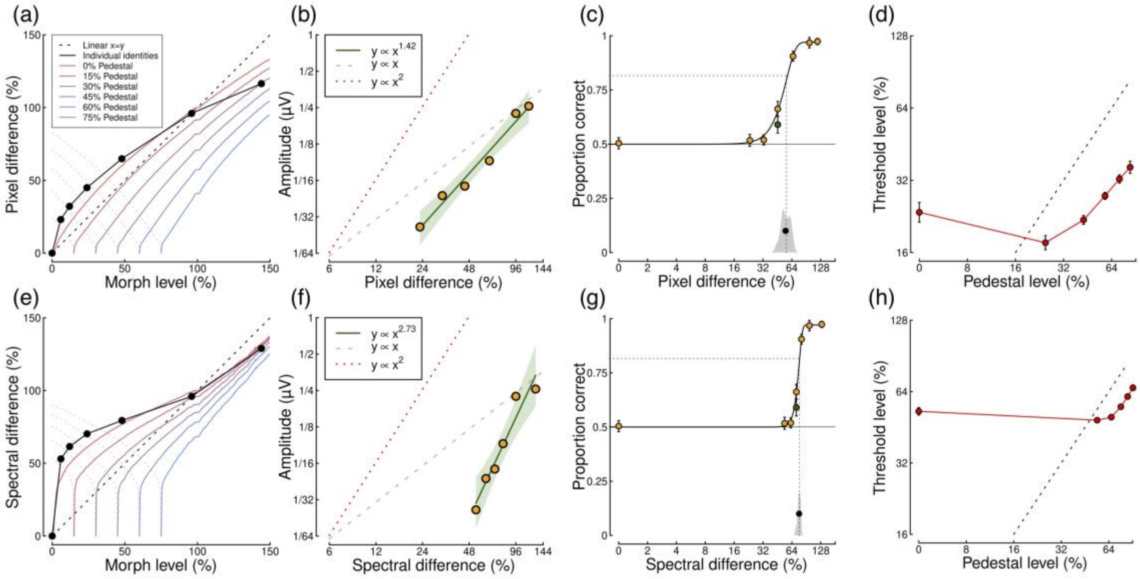

Figure 5: Alternative metrics still support nonlinear processing. Panels (a,e) show how stimuli of different morph levels differ in pixel luminance or Fourier amplitude. Black points show the estimates averaged across the 38 identities used in the first two experiments. Coloured curves show the estimates averaged across the male and female examples used in the discrimination experiment, starting at different pedestal levels. In each case, the values were divided by the difference at 100% (or 96%) morph level and expressed as a percentage, so that the units were comparable to the morph level units used throughout the paper. The oblique dashed line shows the expectation for a linear mapping between units. The remaining panels replot the data from Figures 1e, 2a and 3c using the alternative units, but with the same plotting conventions as described in the relevant figure captions.

To understand how these alternative metrics might influence our conclusions, we re-ran our analyses replacing the (linear) morph levels with the pixel or spectral difference values (rescaled to be in analogous percentage units). Our rationale is that if the nonlinearity in the stimulus is responsible for (some of) the apparently nonlinear processing in the brain, using these alternative units will result in more approximately linear processing. These results are shown in Figure 5, and in Table 1 we report four indices of nonlinearity across the three experiments. Figures 5a,e show how the difference metrics change as a function of morph level. If these were entirely linear all curves would run parallel to the oblique dashed unity line. Clearly there are some substantial deviations, however we note that the very steep portion of the nonlinearity is at small morph levels (<15%) well below detection threshold (see Figure 2a) where neural responses cannot be differentiated from noise (Figure 1d). This means that the main influence of using these alternative units will be determined by the shallower slope evident at higher morph levels.

Table 1: Summary of indices of nonlinearity for different candidate input units. The units summarise the main features of nonlinearity for each experiment, and comprise: the slope of the emotion response function (determined by linear regression on log-log values), the transducer exponent inferred by the slope of the psychometric function (Weibull β/1.3), the amount of facilitation given by the ratio of thresholds between 0% and 15% morph levels of the dipper function, and the slope of the dipper handle (over the four highest pedestal levels). These indices give evidence of nonlinear processing when they deviate from the linear predictions listed in the bottom row.

Input units	SSVEP slope	Weibull β/1.3	Facilitation	Handle
Morph level	0.73	1.78	1.55	0.57
Pixel difference	1.42	3.42	1.34	0.76
Spectral difference	2.73	9.95	1.09	0.90
Linear prediction	1	1	1	0

When using the pixel difference metric, the emotion response function (Figure 5b) and the psychometric function (Figure 5c) are shifted to the right and become steeper. This is because over most of the range of stimulus levels the pixel differences increase with a slope of less than 1 (compare points in Figure 5a with the oblique dashed line). This means that, relative to using the morph level units, a smaller change in the stimulus is required to produce a unit increase in response (or accuracy). The summary indices shown in Table 1 support this – the slope of the emotion response function and the psychometric function both increase relative to those derived using morph level units. The dipper functions also shift to the right and become somewhat steeper, for similar reasons (see Figure 5d). However, the form of the dipper is still apparent, with clear facilitation (a factor of 1.34), and masking in the 'handle' region (with a slope of 0.76). All of these changes become more extreme for the spectral difference metric (Figure 5f-h), yet in all cases there is still evidence of nonlinear processing in the brain. Overall then, our main indices of nonlinearity are changed somewhat by the use of image-based units, but we can still conclude that neural processing of emotion is nonlinear.

We think it relatively unlikely that these low-level image differences are actually used by participants for several reasons. In the psychophysical tasks, participants were explicitly instructed to respond to the emotional content of the stimulus rather than image features such as luminance, spatial frequency and orientation. Viewing the stimuli used in these experiments delivers a compelling subjective experience of changes in emotion, which 'pop out' of the dynamic sequences used in the first two experiments (see Figure 1a). Because we used random identities in this temporal sequence, this will likely confound the low-level changes that might be present within an identity. In addition, we observed strong inversion effects (Eimer & Holmes, 2002; Yin, 1969) in the SSVEP and detection experiments (green points in Figures 1d and 2a). For inverted stimuli, differences in low level image properties remain constant, yet performance and neural responses are both significantly reduced relative to upright stimuli. Finally, making reliable judgements about expression in everyday life is unlikely to be possible using cues such as luminance, which will vary idiosyncratically depending on the situation. It is conceivable that the visual system might use some of the information from lower level features in combination with the expression information, yet our analysis suggests that this would only increase the evidence for nonlinear neural processing.

3.3 Conclusions

Across three experiments using different paradigms and stimuli, we find evidence that facial expression intensity is processed in a nonlinear fashion. These findings are consistent with the idea that relatively weak expressions are most typically experienced in everyday life, and the brain might benefit from increasing sensitivity to subtle changes of expression within this range. We predict that similar nonlinearities might apply along other dimensions of face-space, including facial identity, age, attractiveness, and facial features that communicate character traits such as dominance and trustworthiness. Such nonlinearities would distort the geometry of 'face space' in predictable ways that might be quantified in future studies using the methods developed here.

5. Acknowledgements

We thank Mike Burton and Andy Young for helpful comments on earlier versions of this work. MY was supported by a bursary from the Experimental Psychology Society.

6. References

Adams, W. J., Gray, K. L. H., Garner, M., & Graf, E. W. (2010). High-level face adaptation without awareness. *Psychological Science*, *21*(2), 205–210. https://doi.org/10.1177/0956797609359508

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62. https://doi.org/10.1177/1534582302001001003

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*, *48*(1), 217–237. https://doi.org/10.1152/jn.1982.48.1.217

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*(3), 217–239.

Blair, R. J., Colledge, E., Murray, L., & Mitchell, D. G. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology*, *29*(6), 491–498.

Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision Research*, *39*(2), 257–269.

Braddick, O. J., Wattam-Bell, J., & Atkinson, J. (1986). Orientation-specific cortical responses develop in early infancy. *Nature*, *320*(6063), 617–619. https://doi.org/10.1038/320617a0

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Burr, D., Silva, O., Cicchini, G. M., Banks, M. S., & Morrone, M. C. (2009). Temporal mechanisms of multimodal binding. *Proceedings. Biological Sciences*, *276*(1663), 1761–1769. https://doi.org/10.1098/rspb.2008.1899

Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of concurrent stimuli by population activity in visual cortex. *Neuron*, *64*(6), 931–942. https://doi.org/10.1016/j.neuron.2009.11.004

Butler, A., Oruc, I., Fox, C. J., & Barton, J. J. S. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Research*, *1191*, 116–126. https://doi.org/10.1016/j.brainres.2007.10.101

Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition*, *76*(2), 105–146.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical Perception of Morphed Facial Expressions. *Visual Cognition*, *3*(2), 81–118. https://doi.org/10.1080/713756735

Calder, A. J., Young, A. W., Rowland, D., & Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings. Biological Sciences*, *264*(1383), 919–925. https://doi.org/10.1098/rspb.1997.0127

Campbell, F. W., & Kulikowski, J. J. (1972). The visual evoked potential as a function of contrast of a grating pattern. *The Journal of Physiology*, *222*(2), 345–356.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews. Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Chen, P.-Y., Pan, Y.-C., & Chen, C. C. (2014). The role of upper and lower faces in discrimination happy and sad facial expressions. *IPerception*, *5*(4), 340 [abstract].

Coll, M.-P., Murphy, J., Catmur, C., Bird, G., & Brewer, R. (2019). The importance of stimulus variability when studying face processing using fast periodic visual stimulation: A novel “mixed-emotions” paradigm. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *117*, 182–195. https://doi.org/10.1016/j.cortex.2019.03.006

Dakin, S. C., & Omigie, D. (2009). Psychophysical evidence for a non-linear representation of facial identity. *Vision Research*, *49*(18), 2285–2296. https://doi.org/10.1016/j.visres.2009.06.016

Dzhelyova, M., Jacques, C., & Rossion, B. (2017). At a Single Glance: Fast Periodic Visual Stimulation Uncovers the Spatio-Temporal Dynamics of Brief Facial Expression Changes in the Human Brain. *Cerebral Cortex (New York, N.Y.: 1991)*, *27*(8), 4106–4123. https://doi.org/10.1093/cercor/bhw223

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *Neuroreport*, *13*(4), 427–431.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227–240. https://doi.org/10.1016/0010-0277(92)90002-Y

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research*, *1127*(1), 80–89. https://doi.org/10.1016/j.brainres.2006.09.104

Georgeson, M. A., Yates, T. A., & Schofield, A. J. (2008). Discriminating depth in corrugated stereo surfaces: facilitation by a pedestal is explained by removal of uncertainty. *Vision Research*, *48*(21), 2321–2328. https://doi.org/10.1016/j.visres.2008.07.009

Gori, M., Mazzilli, G., Sandini, G., & Burr, D. (2011). Cross-Sensory Facilitation Reveals Neural Interactions between Visual and Tactile Motion in Humans. *Frontiers in Psychology*, *2*, 55. https://doi.org/10.3389/fpsyg.2011.00055

Harris, R. J., Young, A. W., & Andrews, T. J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proceedings of the National Academy of Sciences*, *109*(51), 21164–21169. https://doi.org/10.1073/pnas.1212207110

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*(2), 181–197.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*, 241–257. https://doi.org/10.1023/A:1024952730333

Juricevic, I., & Webster, M. A. (2012). Selectivity of face aftereffects for expressions and anti-expressions. *Frontiers in Psychology*, *3*, 4. https://doi.org/10.3389/fpsyg.2012.00004

Kingdom, F. A. A. (2016). Fixed versus variable internal noise in contrast transduction: The significance of Whittle's data. *Vision Research*, *128*, 1–5. https://doi.org/10.1016/j.visres.2016.09.004

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, *70*(12), 1458–1471.

Leleu, A., Dzhelyova, M., Rossion, B., Brochard, R., Durand, K., Schaal, B., & Baudouin, J.-Y. (2018). Tuning functions for automatic detection of brief changes of facial expression in the human brain. *NeuroImage*, *179*, 235–251. https://doi.org/10.1016/j.neuroimage.2018.06.048

Linares, D., & López-Moliner, J. (2016). quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. *The R Journal*, *8*(1), 122. https://doi.org/10.32614/RJ-2016-008

Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2014). An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia*, *52*, 57–72. https://doi.org/10.1016/j.neuropsychologia.2013.10.022

Marneweck, M., Loftus, A., & Hammond, G. (2013). Psychophysical Measures of Sensitivity to Facial Expression of Emotion. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00063

Meese, T. S., & Summers, R. J. (2012). Theory and data for area summation of contrast with and without uncertainty: Evidence for a noisy energy model. *Journal of Vision*, *12*(11), 9–9. https://doi.org/10.1167/12.11.9

Meese, Tim S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision*, *6*(11), 1224–1243. https://doi.org/10.1167/6.11.7

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A “dipper” function for texture discrimination based on orientation variance. *Journal of Vision*, *8*(11), 9.1-8. https://doi.org/10.1167/8.11.9

Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research*, *21*(2), 215–223. https://doi.org/10.1016/0042-6989(81)90115-2

Nachmias, J., & Sansbury, R. V. (1974). Letter: Grating contrast: discrimination may be better than detection. *Vision Research*, *14*(10), 1039–1042.

Nelson, P. C., & Carney, L. H. (2006). Cues for masked amplitude-modulation detection. *The Journal of the Acoustical Society of America*, *120*(2), 978–990.

Öhman, A. (2002). Automaticity and the Amygdala: Nonconscious Responses to Emotional Faces. *Current Directions in Psychological Science*, *11*(2), 62–66. https://doi.org/10.1111/1467-8721.00169

Ohzawa, I., Sclar, G., & Freeman, R. D. (1982). Contrast gain control in the cat visual cortex. *Nature*, *298*(5871), 266–268.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Orgeta, V., & Phillips, L. H. (2008). Effects of age and emotional intensity on the recognition of facial emotion. *Experimental Aging Research*, *34*(1), 63–79. https://doi.org/10.1080/03610730701762047

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, Optics and Image Science*, *2*(9), 1508–1532.

Raab, D. H., Osman, E., & Rich, E. (1963). Intensity Discrimination, the “Pedestal” Effect, and “Negative Masking” with White-Noise Stimuli. *The Journal of the Acoustical Society of America*, *35*(7), 1053–1053. https://doi.org/10.1121/1.1918653

Reynaud, A., Barthélemy, F. V., Masson, G. S., & Chavane, F. (2007). Input-output transformation in the visuo-oculomotor loop: comparison of real-time optical imaging recordings in V1 to ocular following responses upon center-surround stimulation. *Archives Italiennes De Biologie*, *145*(3–4), 251–262.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139–253. https://doi.org/10.1080/13506285.2013.772929

Rossion, B., Prieto, E. A., Boremanse, A., Kuefner, D., & Van Belle, G. (2012). A steady-state visual evoked potential approach to individual face perception: effect of inversion, contrast-reversal and temporal dynamics. *NeuroImage*, *63*(3), 1585–1600. https://doi.org/10.1016/j.neuroimage.2012.08.033

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience*, *8*(1), 107–113. https://doi.org/10.1038/nn1370

Russell, J. A., & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology*, *22*(1), 97–102. https://doi.org/10.1037/0012-1649.22.1.97

Tanaka, J., Giles, M., Kremen, S., & Simon, V. (1998). Mapping attractor fields in face space: the atypicality bias in face recognition. *Cognition*, *68*(3), 199–220.

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

Tsai, J. J., Wade, A. R., & Norcia, A. M. (2012). Dynamics of normalization underlying masking in human visual cortex. *The Journal of Neuroscience*, *32*(8), 2783–2789. https://doi.org/10.1523/JNEUROSCI.4485-11.2012

Tyler, C. W., & Chen, C. C. (2000). Signal detection theory in the 2AFC paradigm: attention, channel uncertainty and probability summation. *Vision Research*, *40*(22), 3121–3144.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(2), 161–204.

Wallis, S. A., Baker, D. H., Meese, T. S., & Georgeson, M. A. (2013). The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision. *Vision Research*, *76*, 1–10. https://doi.org/10.1016/j.visres.2012.09.019

Watt, R. J., & Morgan, M. J. (1983). The recognition and representation of edge blur: evidence for spatial primitives in human vision. *Vision Research*, *23*(12), 1465–1477.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*(6982), 557–561. https://doi.org/10.1038/nature02420

Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology*, *92*(3), 1830–1839. https://doi.org/10.1152/jn.00155.2004

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145. https://doi.org/10.1037/h0027474

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*(3), 271–313.

780