Evolution and Human Behavior xxx (xxxx) xxx

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens

Signal value of stress behaviour

Jamie Whitehouse[a,*], Sophie J. Milward[b], Matthew O. Parker[c], Eithne Kavanagh[a], Bridget M. Waller[a]

[a] Department of Psychology, Nottingham Trent University, Nottingham, UK
[b] Department of Psychology, University of Portsmouth, Portsmouth, UK
[c] School of Pharmacy and Biomedical Sciences, University of Portsmouth, Portsmouth, UK

ARTICLE INFO

Keywords:
Stress
Individual differences
Displacement behaviour
Facial expressivity

ABSTRACT

Physiological and psychological stress are accompanied by nonverbal behaviour across a wide range of species. The function of this 'stress behaviour' is not well understood but is often assumed to be read by others as a cue to stress. Displaying signs of weakness is, however, difficult to understand from an evolutionary perspective and therefore further investigation into why these behaviours exist is needed. Here, we test whether displacement behaviours (i.e., those known to be associated with stress) are reliable indicators of stress in humans. To do this, we presented raters ($N = 133$) with videos of individuals ($N = 31$) undergoing a stress-inducting task. Self-directed displacement behaviours and self-reported stress were both associated with stress ratings given by raters. Therefore, such behaviours can provide reliable information to others and can be considered communicative. Individuals producing more nonverbal stress behaviour were rated as more likeable by raters (perhaps presenting as more honest signallers), indicating a benefit and potential adaptive function of displaying stress. Raters also differed in their accuracy in detecting stress from nonverbal cues. Findings suggest that the accuracy with which individuals were able to detect stress was linked to the number of social connections they reported to have. However, this association was non-linear, with individuals who were most and least accurate reporting the least network connections. This could indicate that the ability to read behaviour is associated with an ability to form and maintain social networks.

1. Introduction

The experience of stress is often paired with visual cues across a wide range of species (Troisi, 1999). These include self-directed behaviours such as scratching, face touching and lip-biting (Troisi, 2002) as well as some stress-specific facial movements (Giannakakis et al., 2017; Mayo & Heilig, 2019). There is a body of behavioural, physiological and pharmacological evidence demonstrating that the experience of stress is reliably associated with the production of these behaviours (Maestripieri, Schino, Aureli, & Troisi, 1992; Mohiyeddini, Bauer, & Semple, 2015), and their study has gained significant attention in the clinical and health sciences (Mayo & Heilig, 2019; Troisi, 1999) due to their important applied benefits in psychiatry. However, science has somewhat neglected questions concerning why these behaviours evolved and what adaptive benefits they could provide to a stressed individual and those around them (Tinbergen, 1952). Thus, why stress behaviours actually exist remains a mystery. The emergence of these behaviours is not immediately obvious, as displaying overt visual signs of weakness is difficult to understand from an evolutionary perspective. Producing behaviours so strongly associated with stress could provide opportunities for an individual to be taken advantage of by others and an adaptive strategy should be to conceal stress and other weakened states. But at least for stress, this does not seem to be the case. This could reflect the cooperative nature of humans (Tomasello, 2010) and that this risk of competition from others simply does not exist. Or, it could be that the benefits gained from displaying stress to others outweigh the risks of competition, such as providing key opportunities to elicit empathy and help from others (Dezecache, Jacob, & Grèzes, 2015); social benefits which could act as a strong selection pressure. As a comparison, crying (a signal of negative affect) is known to elicit or enhance shared emotional experience (Preston & de Waal, 2002). However, all this this assumes that stress behaviours are salient to others. It may be that these behaviours are not actually perceived as associated with stress in real-world social interactions, for example, because of their significant

* Corresponding author.
E-mail address: jamie.whitehouse@ntu.ac.uk (J. Whitehouse).

https://doi.org/10.1016/j.evolhumbehav.2022.04.001
Received 4 May 2021; Received in revised form 23 February 2022; Accepted 5 April 2022

overlap with other behaviours. The scratching and self-grooming behaviours often associated with stress also have the more basic function of removing irritations from the skin (Maestripieri et al., 1992). Such alternate and non-affective functions could make any information transfer about stress noisy and unreliable if they are also produced readily in situations where stress is not present. All this currently remains unknown, however, as the accuracy in which humans can perceive the stress of others (and what factors determine this accuracy) is yet to be quantified in detail.

Stress behaviours are frequently referred to as *displacement behaviours* (Mohiyeddini & Semple, 2013), a term coined by early ethologists who thought their emergence was as a consequence of *displaced* energy (Troisi, 2002). It was proposed that during a motivational/internal conflict (for example, when there is both motivation for conflict and affiliation simultaneously, McFarland, 1966), a third behaviour often irrelevant to the context is often produced, causing a distraction from the negative stimuli or acting as a 'sensory cut-off' (Chance, 1962). In humans, these behaviours appear to mainly manifest through self-directed 'comfort' behaviours such as self-grooming face-touching, head scratching and through behaviours through the iterative manipulation of objects such as fumbling with jewellery, and chewing on pens (Troisi, 1999). These behaviours may have proximate function to regulate the experience of stress (Mohiyeddini, Bauer, & Semple, 2013), as individuals who produce more stress-associated behaviour seem to recover from a stressful event quicker; measured through lower self-reported stress (Mohiyeddini & Semple, 2013) and through lower heart rate post stressful event (Pico-Alfonso et al., 2007). In non-human primates, there is also strong pharmacological evidence linking these behaviours to stress, and increased rates of self-scratching in monkeys are positively associated with the administration of anxiety-inducing drugs (and negatively associated with anxiety-relieving drugs) (Troisi, 2002).

The observable link between stress and behaviour suggests a communicative function of these behaviours, either as a signal that has been specifically selected and evolved, or as a phenomena occurring as a by-product of other functions (i.e., a cue: Scott-Phillips, Blythe, Gardner, & West, 2012). Although stress associated behaviours have been proposed to be communicative in the past (Bradshaw, 1993; Maestripieri et al., 1992), these inferences have been made largely from quantifying the relationship between individuals' experienced stress and their behaviour. However, to understand if the emergence of these behaviours was driven (or at least, in some part driven) by a signal function, a shift in focus from the producer of these behaviours to the psychology and behaviour of the observer (or receiver; Guilford & Dawkins, 1991) is necessary. This perspective has already been attempted with species of non-human primate (macaques: Whitehouse, Micheletta, & Waller, 2017), who were found to interact differently (i.e. more affiliatively) around individuals displaying stress behaviours, appearing to respond to these behaviours in a meaningful and adaptive way. As we share many similarities in the way stress is manifested in behaviour compared with non-human primates (many species of which have also been reported to produce self-directed behaviour during periods of stress, Maestripieri et al., 1992; Troisi, 2002), it is possible that we also share function, and that producing these behaviours affords us comparable social benefits.

People can vary in both the extent to which they produce signals (Kanai & Rees, 2011; Mayo & Heilig, 2019), and their ability to accurately read and process them (Duesenberg et al., 2016). Whether these individual differences represent noise, or if there is an adaptive reason for people to behave differently, is an unknown but interesting issue. Communicative complexity is known to be linked to social complexity (Freeberg, Dunbar, & Ord, 2012) and it is thought that those species with more complex communication are consequently able to navigate a more complex social environment. We could also hypothesise a similar relationship at an individual level and expect that someone's communicative skills are then positively associated with their social environment and their ability to maintain bigger social networks. Testing such a hypothesis with stress behaviours could be very useful. If producing or processing stress behaviours is linked to the size of an individual's social network, this would provide further evidence that stress behaviours function within a social environment and that they have an evolved signal function. Additionally, looking at the relationship between an individual's ability to display and detect stress and their social environment may even help us explain why we can observe individual variability within expressivity in general.

In order to assess for a potential signal function of stress, or more simply, if people are able to recognise the stress in others by observing their behaviour, we designed an experiment. A group of participants (termed 'actors' here on) were exposed to a stressful task whilst videotaped, and several indices of stress, including self-report, behavioural, and physiological measures, were taken. We presented these videos (without an explanation about the context in which they were obtained) to a larger pool of participants (termed 'raters' here on) who were required to give these videos ratings (on the actors' experience of stress and their likeability) and provide some details regarding their own social network. Our aim was to explore the relationship between stress, behaviour and its perception. Although a largely exploratory approach was taken, we had the following predictions: First, we hypothesised that the ratings given by the raters would align with the measures of stress taken from the actors, and that the most reliable indicator of stress in others would be displacement behaviour (i.e. stress behaviours). Second, we hypothesised that those who appear more stressed would also be judged as more likeable, a finding which would further suggest an adaptive benefit to the communication of stress (i.e., relationship building). Finally, we assessed how individual variation in stress-reading skills reflected social network size, with the idea these skills are directly linked to an individual's ability to navigate and develop their social environment. Here we predicted that better stress-reading skills would afford a larger social network.

2. Methods

2.1. Participants (Actors)

To collect the stimuli for our subsequent experiments, thirty-one participants were recruited (23 female, mean age = 28.5 ± 10.5). The task presented to these participants was in 5 parts: 1) pre-task questionnaires, 2) pre-task salivary sample (see Section 2.2.3 for saliva sampling method), 3) stress-inducing task (Trier-social stress test), 4) post-task salivary sample, 5) post-task questionnaire. Participants were recruited and compensated in person. The collection of these data received ethical approval from the University of Portsmouth science faculty ethics committee (Reference Number: SFEC 2019-001).

2.2. Experimental procedure (Actors)

2.2.1. Trier social stress test

In order to induce stress in our participants, we invited them to engage with the Trier-Social Stress Test (TSST, see Kirschbaum, Pirke, & Hellhammer, 1993). This task is made up of 5 phases. 1) A c*ontrol* phase where participants are asked demographic questions and some simple questions about themselves (e.g. 'what is your name?', 'what is your favourite food?', 'do you have any pets?). 2) a *preparation* phase where participants must prepare for a clinical psychologist job interview, including a 3-minute speech about why they are an ideal candidate. 3) Test phase 1, where participants are required to give their speech. 4) Test phase 2, where participants are given a series of purposefully difficult job interview questions (e.g. 'What do you consider to be the main advances in the area of Clinical Psychology in the last 20 years?'). 5) Test phase 3, where participants would engage in a mental arithmetic task ("Please count backwards from the number 1022 in 13's as fast and as accurately as you can until I ask you to stop."). Throughout the entire process, the participants full body was filmed on a blank background.

Afterwards the participant was reassured and debriefed about the aims of this task.

2.2.2. Questionnaires

Participants were presented with questions pre- and post-TSST. Pre-task, they were asked "How stressed do you feel?" on a 1–100 scale. Post-task they were asked the same question as above, but where additionally presented with the state-trait anxiety questionnaire (STAI, Spielberger, 1983), which aims to provide a quantitative measure of both the participants state and trait anxiety.

2.2.3. Cortisol

Saliva samples were collected by the 'passive drool' method. Briefly, participants were asked to tilt their head forward and activate their saliva glands with their tongue until a pool of saliva accumulated in the mouth. They were then asked to gently spit the saliva into a 15 ml falcon tube (Sigma, UK). Samples were immediately put on ice prior to being transported to the laboratory for processing. Any samples that contained visible contamination with blood were discarded. All saliva samples were thawed, vortexed (2-min) and centrifuged at 1500 $\times g$ for 15-min and the top layer (containing cortisol) was removed and stored at −20 °C until assay.

The salivary cortisol concentration was assayed using a Salivary Cortisol Enzyme Immunoassay (ELISA) according to the manufacturer's protocol (Salimetrics 1-3002 96-well kit, Stratech, UK). All samples were run in duplicate (inter-assay Coefficient of Variation 3.2%). Average optical density (OD) was calculated for all duplicate wells. We subtracted the average non-specific binding (NSB) wells from all OD values. We calculated the percent bound in each sample by dividing by the average OD of the 'zero' (Bo) wells (B/Bo). Finally, we determined the concentration of cortisol in each sample using a 4-paramter non-linear regression curve fit using an online free analysis programme (https://elisaanalysis.com).

2.3. Video stimuli preparation

From each video collected during the TSST, we prepared an 80 s video clip to be used as stimuli for a subsequent rating experiment. These video clips were extracted from the onset of the arithmetic task. This video segment was chosen for multiple reasons; first, as it is the final task of the TSST, is it the closest (in time) to their post-test stress rating (and thus, we could expect this rating to be more indicative of the stress experienced during the video clip). Second, it is the most standardised and least ambiguous task of the TSST (e.g. there can be, and were, bigger individual differences in how people approach an interview), and therefore we believe the arithmetic task is the most appropriate to directly behaviours between actors. Thirty-one videos were exported as an .mp4 (1080p resolution, 24FPS) using Adobe Premier Pro (http://www.adobe.com/). Videos were exported without audio.

2.4. Participants (Raters)

We recruited 133 participants (94 female) to act as raters. All raters were aged between 35 and 45 (mean: 39 ± 3.2) years old. As the composition of an individual's social network tends to vary across age when considering both the quantity and stability of social connections, we recruited individuals from within a small age-range (within 10 years), and at an age in where social networks are thought to be most stable (around 40 years of age; Morgan, 1988).

Raters were recruited and compensated (through the online platform prolific (http://www.prolific.co, at a rate of £6.50/h) where they were redirected to an online based experiment developed and hosted by Qualtrics (http://www.qualtrics.com). There were 4 parts to this experiment: (1) consent and demographic questionnaire, (2) the social network index questionnaire (Cohen, Doyle, Skoner, Rabin, & Gwaltney, 1997), (3) Berkeley expressivity questionnaire (Gross & John, 1997) and finally (4) a video-rating exercise where participants were required to watch and subsequently rate 10 videos. Participants took between 25 and 35 min to complete the experiment. The collection of these data received ethical approval from the University of Portsmouth science faculty ethics committee (Reference Number: Reference Number: SFEC 2020-042).

2.5. Experimental procedure (Raters)

2.5.1. Questionnaires

Here, we were interested in two measures of individual difference in our raters; differences in social network size, and differences in emotional expression. In order to gain a quantitative measure of each of these, we presented our subjects with two pre-established questionnaires, the Social Network Index (SNI; Cohen et al., 1997) and the Berkeley Expressivity Questionnaire (BEQ; Gross & John, 1997). Prior to these, participants were also asked for their age and gender.

The SNI questionnaire, developed by Cohen et al. (1997) assesses an individual's social connection based on 12 social relationship types (e.g. partner, friend, neighbour, work colleague). Participants are asked questions such as, *"How many of your children do you see or talk to at least once in a usual two-week period (including phone calls, video chat, or instant messaging)",* receiving +1 to their social network size for every individual whom they report engaging with at least once in a usual 2-week period.

The BEQ questionnaire, developed by Gross and John (1997) assess three facets of an individual's emotional expressivity: negative expressivity, positive expressivity and impulse strength as well as providing a score of overall expressivity. The BEQ is a 16-part questionnaire, with each response on a 7-point Likert scale ranging from *disagree strongly* to *agree strongly*. Each question is associated to a specific facet, e.g. negative expressivity; *"Whenever I feel negative emotions, people can easily see exactly what I am feeling.",* positive expressivity; *"When I'm happy, my feelings show.",* impulse strength; *"I experience my emotions very strongly.*". A score from 1 to 7 is given to each participant for each facet, which is a mean of the scores provided for all the facet-associated questions. Overall expressivity is calculated from the mean score of the three facets.

More details about the specific questions included in the questionnaires can be found in the supplementary materials (SM1).

2.5.2. Video rating task

Participants were then shown a random sample of 10 stimuli videos (out of the possible 31) and were prompted to answer the following three questions after the video had finished; 1) *"How stressed does this person (in the video) look?"* 2) *"How confident are you in this judgment?"* and 3) *"As a first impression, how much do you like this person?"*. Each question was presented on a sliding scale from 0 to 100. Videos with their following questions were presented one-by-one, and to encourage attention to the videos, participants could not progress to the next rating task until the 80 s (the length of the video) had passed. 133 participants viewed 10 random videos each out of a potential 31, and from these we collected 1326 individual ratings (in 4 occurrences the participant failed to make a rating). This process resulted in an average of 42.7 (± 1.38SD) ratings per video.

2.6. Video coding

2.6.1. Global behaviours (ESCI)

In order to compare the actor's behaviour in the video to the ratings they were given, we conducted behavioural coding on all rated videos according to the Troisi (1999) modified Ethological Coding System for Interviews (ECSI). This coding system includes 37 different behavioural patterns, mainly consisting of hand, head and face movements, which contribute to seven behavioural categories: Affiliative behaviours, Submissive behaviours, Prosocial behaviours, Flight behaviours, Assertive behaviours, Displacement behaviours and Relaxed behaviours. As an example, behaviours making up *Displacement behaviours* include groom, hand-to-face, hand-to-mouth, scratching, yawning, fumbling, twisting the mouth, licking lips and biting lips, and the summed duration of these would provide a Displacement score. From this, we also calculated these behavioural categories as a proportion to all behaviours observed (e.g. 20% of all behaviours observed are Affiliative). In order to retain statistically independent behavioural categories, we removed Assertive, Flight and Prosocial categories as these had behaviours which were scored over multiple categories. A full list and description of each of the coded behaviours can be found in the supplementary materials (SM2).

All videos were coded using BORIS (Behavioural Observation Research Interactive Software, http://www.boris.unito.it, Friard & Gamba, 2016). Inter-rater reliability was conducted between JW and EK on 15% (5) videos. In this process, we compared the durations reported during each behavioural occurrence. There was a significant agreement in the total time each participants engaged in each behaviour $r(80) = 0.71$, $p < 0.001$ (Pearson's correlation).

2.6.2. Facial movements (FACS: Facial action coding system)

For all rated videos, facial movements were coded in further detail using the Facial Action Coding System (FACS, Ekman & Friesen, 1978) by two certified FACS coders JW and EK. In the FACS coding scheme, the onset and offset of individual facial muscle movements (Action Units; AUs) are recorded. FACS provides an objective measure of facial movements and allows for a quantitative measure of expressivity without any underlying assumptions about the emotions of the actor. The FACS system recognises 28 movements connected to individual muscles such as inner brow raiser (AU1), nose wrinkle (AU9) and lip-corner puller (AU12). It also recognised an addition 28 movements that may involve the co-occurrence of multiple muscles such as jaw clencher (Action Descriptor 31), cheek suck (AD35) and lip wipe (AD37). Any lower face action units which could not be disassociated from movements due to speech were not coded (e.g. lip pressor, AU24). In addition, FACS normally includes head and eye movements (e.g. Look up, Head left), however, these were not coded for the purposes of this study. A full list of all AUs coded and their descriptions can be found in the supplementary materials (SM2).

From the FACS data we calculated four measures of expressivity for each actor's video: 1) total AU frequency (the frequency of AUs), 2) total AU duration (the combined duration of activated AUs), 3) AU repertoire (the number of unique AU's produced), and 4) AU diversity. AU diversity was calculated following the methods described by Scheider, Liebal, Oña, Burrows, and Waller (2014), and is a measure which incorporates both AU repertoire and duration and informs us about how evenly facial expressions are observed. AU diversity score is maximised when all AU types in an individual's repertoire are evenly observed. See Scheider et al. (2014) for further details.

All videos were coded using BORIS (Behavioural Observation Research Interactive Software, http://www.boris.unito.it, Friard & Gamba, 2016). Inter-rater reliability was conducted on all FACS coding between JW and EK on 15% (5) of the videos, and significant agreement was found. For frequency of behaviour, coders had an average agreement of 0.72 (using Wexler's agreement, Ekman & Friesen, 1978) which is considered good for FACS data. For duration of behaviour, there was significant correlation between the data of the two coders, $r(257) = 0.61$, $p < 0.001$ (Pearson's correlation).

2.7. Statistical analysis

2.7.1. Part A. Assessing the relationship between stress rating and behaviour

The main goal of our first analysis was to assess which of our variables collected from the actor's self-report and spontaneous behaviour most accurately reflected the stress-ratings provided by the raters. This analysis followed a correlation design, where each of our considered variables (Table 1) were compared through a series of Pearson's correlations. Each data point in this analysis was an actor (a stimulus video). From this, we could broadly infer which variables were most strongly driving the stress ratings, but also, which of our other variables co-varied. Following recommendations from Bender and Lange (2001) for exploratory analyses, we prioritised reducing the risk of type 2 over type 1 error and so did not control for multiple comparisons.

2.7.2. Part B. Exploring individual differences in stress rating accuracy

The main goal of our second analysis was to quantify the accuracy of each individual rater (i.e. how skilled individual raters were at using behavioural queues to inform their stress-ratings), and to assess what individual differences predict this accuracy. To do this, we first built a generalised linear model (using function *glm* in r), which assessed how the proportion of displacement behaviour produced by our actors predicted the average stress-rating of each video. Here, we decided to use displacement behaviour as our most accurate marker of stress in our actors as this most closely aligns with previous literature, and heavily influenced our raters in this study. Based on this model, we extracted the predicted stress-ratings for each video (by using the *predict* function on the original data). For each unique stress-rating in this study ($n = 1326$), we then calculated the absolute error between the stress-rating scored by

Table 1
The 16 variables considered in our analysis.

Variables ($n = 31$)	A/R	Description
Stress rating	R	The mean stress rating scored for that actor (values from 1 to 100, calculated from an average of 42.7 ± 1.38 SD raters).
Likeability	R	The mean rating for likeability for that actor (values from 1 to 100).
Self-reported stress	A	The self-reported rating of stress given by the actor, directed following the TSST (values from 1 to 100).
Trait anxiety (29)	A	The score for trait anxiety extracted from the state-trait anxiety inventory.
Cortisol post-task (26)	A	The level of salivary cortisol post-TSST (values in nanomoles per litre, with an average of 25.97nmol/L ± 28.23SD).
ΔCortisol (26)	A	The difference in salivary cortisol levels pre-and post TSST (post-test levels – pre-test levels, average of 2.82 ± 16.01SD).
Affiliative behaviours (duration)	A	The total duration of affiliative behaviours coded; behaviours taken from the modified ethogram for clinical interviews (average of 20.1s ± 16.3).
Affiliative behaviours (proportion)	A	The proportion of affiliative behaviours observed compared to all behaviours (affiliation behaviour duration / all behaviour duration, average of 0.12 ± 0.09).
Displacement behaviours (duration)	A	The total duration of affiliative behaviours coded (average of 24.3s ± 29.2).
Displacement behaviour (proportion)	A	The proportion of displacement behaviours (displacement behaviour duration / all behaviour duration, average of 0.13 ± 0.14).
Submissive behaviours (duration)	A	The total duration of submissive behaviours coded (average of 3.5s ± 4.9).
Submissive behaviours (proportion)	A	The proportion of submissive behaviours (submissive behaviour duration / all behaviour duration, average of 0.02 ± 0.03).
AU frequency	A	The number of facial action units observed (average of 34.2 ± 21.2).
AU duration	A	The total duration of active facial action units observed (98.8s ± 68.0).
AU repertoire	A	The number of unique AU's (average of 11.2 ± 46.4).
AU diversity	A	The AU diversity score (See Scheider et al., 2014 for further details).

A/R = Variable extracted from the Actor (A) or the Rater (R). Numbers in brackets represent the n when missing data is not available for all individuals.

our raters, and the stress-rating predicted by the model. From this, we were able to calculate a mean error for each rater ($n = 133$), as well as the standard deviation of these error values. This gave us, what we call hereafter, an *average error,* and a *variance of error* for each rater.

To assess how the ability to accurately detect stress informs our measures of individual differences in our raters, we computed a series of linear models. Firstly, we built a model to assess how the average error and variance of error ratings (and the interaction between these) predicts the size of the raters' social network (SNI score). The average error and variance of error provide datapoints for accuracy in detecting stress. Second, we built models to assess how these same variables predict the raters' emotional expressivity (taken from the BEQ scoring; positive facet, negative facet, impulse strength, and overall expressivity score). We rebuilt these models for both male and female raters; it was decided not to include gender as an additional interaction term in the model to minimise the complexity of the model at this sample size. Finally, to assess how the confidence in stress rating influences error, we built a generalised linear mixed model (function *lmer* in r) on all rating data, looking at the influence of confidence on average error; here we included the rater and actor as random effects to account for non-independent data.

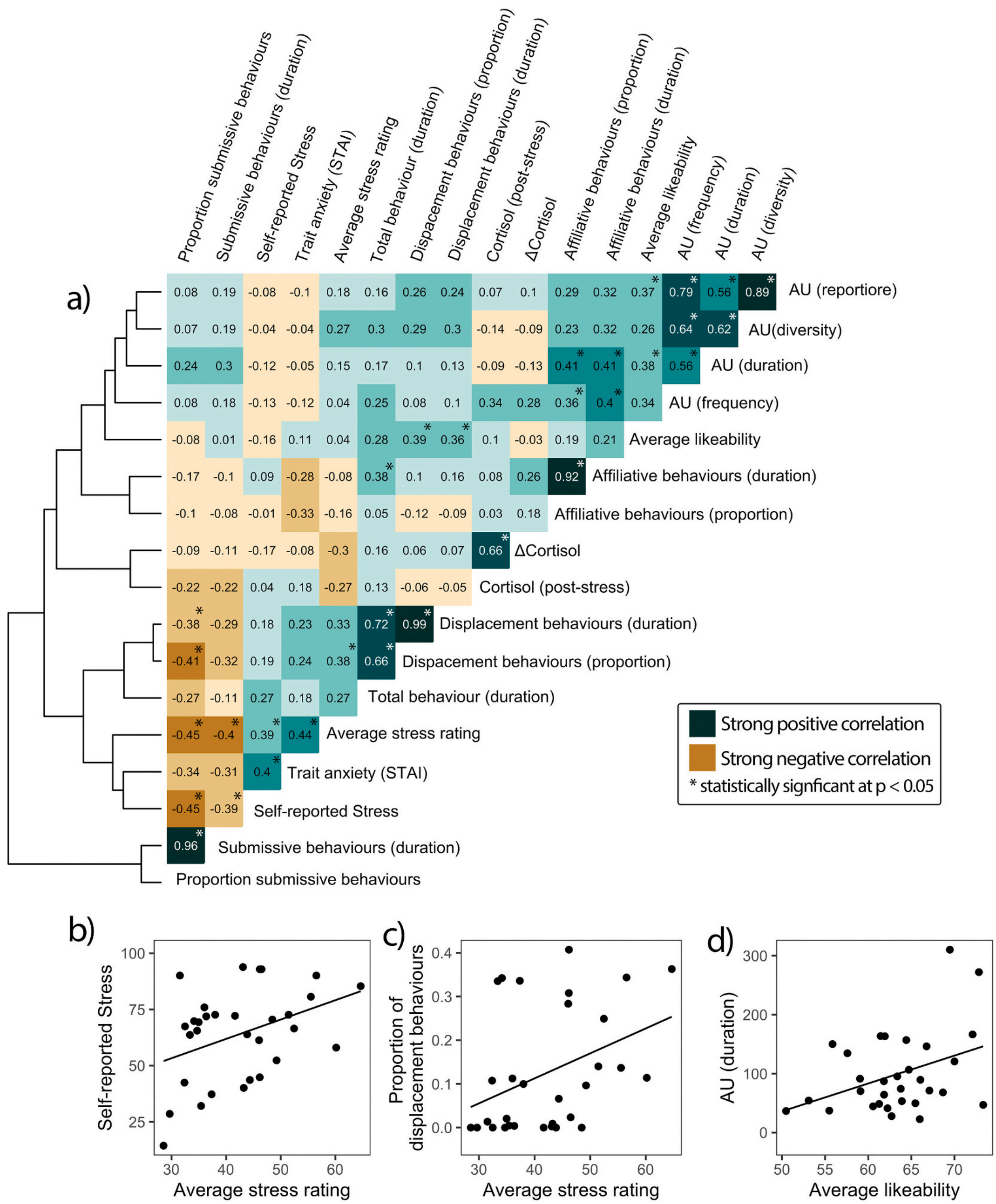

Fig. 1. a) Heatmap demonstrating correlations between all variables. Darker squares represent a stronger correlation (i.e. an r value further away from 0) and lighter squares represent a weaker association (i.e. an r value closer to 0). Pearson's correlations were conducted to test for significant relationships between variables. Variables are clustered into items which are most associated; clusters which are also represented by a dendrogram on the left of the heatmap. b) Significant association between self-reported stress and average stress rating. Each point represents an actor (individual video; $N = 31$). Both X and Y-axis are on a continuous scale of 1–100. c) Significant association between proportion of displacement behaviours and average stress rating. Each point represents an actor (individual video; $N = 31$). The X-axis is on a continuous scale of 1–100, Yaxis represents a proportion between 0 and 1. d) Significant association between likability and AU duration. Each point represents an actor (individual video; $N = 31$). The X-axis is on a continuous scale of 1–100, Y-axis represents the duration of all AU production in seconds.

3. Results

3.1. Part A

The self-reported stress of the actors was positively associated with the average rating of stress (Fig. 1b, $r(29) = 0.39$, $p = 0.029$). The average rating of stress was also positively associated with the proportion of displacement behaviours observed (Fig. 1c, $r(29) = 0.38$, $p = 0.036$) and trait anxiety of the actor ($r(27) = 0.44$, $p = 0.016$), and negatively associated with the proportion and duration of submissive behaviours observed ($r(29) = -0.45$, $p = 0.011$ and $r(29) = -0.40$, $p = 0.026$ respectively). How likeable the actor was judged to be, was positively associated with the proportion of displacement behaviours produced ($r(29) = 0.39$, $p = 0.031$, as well as multiple measures of facial expressivity (Fig. 1d, AU duration, $r(29) = 0.38$, $p = 0.033$; and AU repertoire, $r(29) = 0.37$, $p = 0.038$), but not with the overall amount of behaviour produced. Our hormonal measure of stress via salivary cortisol was not associated with either self-reported or, average stress rating, nor with any other variable considered. Please see Fig. 1a for an overview of all comparisons. From our data, we suspected that the relationship between self-reported stress and average rating of stress may be mediated by displacement behaviours, however a bootstrap analysis testing this mediation model did not show this ($\beta = 0.0273$, 95%CI = −0.031-0.10, $p = 0.404$). Please see section SM4 in the supplementary materials for more details about this.

As there was a sex imbalance in our samples (many more females), we ran additional analyses to confirm that these above results were not driven by our female participants. Overall, we found very little difference in the behaviours and ratings observed in both males and females. The only notable differences were that female actors in general were more likeable than male actors (Wilcoxon signed ranks test: $w = 42$, $p = 0.023$, Fig. SM5c), and that this likeability rating was driven by ratings by female raters (GLMM: $\beta = 9.175$, $SE = 3.507$, $t = 2.616$, $p = 0.0097$, Fig. SM5g). This, however, did not influence any of the relationships found with the other variables and therefore we present our full sample here. Please see the supplementary materials for a more detailed exploration of these sex difference data (SM5).

These findings suggest that stress can be accurately detected by others, and that this detection of stress could be driven by reading of the actors' behaviour. The perception of how likeable a person is perceived was influenced by the presence of these 'stress' behaviours (i.e. displacement behaviours) specifically, and not the overall duration of behaviour in general.

3.2. Part B

Our *glm* confirmed that the proportion of displacement behaviour observed in the videos significantly predicts their average rating of stress ($ß = 24.996$, $SE = 11.366$ $t = 2.199$, $p = 0.036$). From this model, an error for each individual stress-rating was calculated, followed by an *average error* and *variance of error* for each rater (see methods for further details).

Our models assessing the effect of average error and variance of error on social network size (SNI) found a significant interaction between average error and variance of error ($ß = -0.049$, $SE = 0.022$, $t = -2.175$, $p = 0.032$). There were no main effects of average error and variance of error on their own ($ß = 0.484$, $SE = 0.257$, $t = 1.883$, $p = 0.062$; $ß = 1.052$, $SE = 0.541$, $t = 1.945$, $p = 0.054$ respectively). After a visual inspection of the data, we decided to also fit a quadratic term to this interaction, which produced a slightly better fitting model and stronger effect ($ß = -15.72$, $SE = 6.74$, $t = -2.334$, $p = 0.02$). This indicates that an individual's stress recognition skill (i.e., the ability to rate stress both accurately and consistently) was predictive of their social network size. This relationship between stress recognition and social network size appears non-linear (Fig. 2a). Individuals who demonstrated both high average error in their ratings and a higher variance in their error scores (i.e, were most inaccurate in their ratings) reported a smaller social network size. Similarly, those who demonstrated both low average error

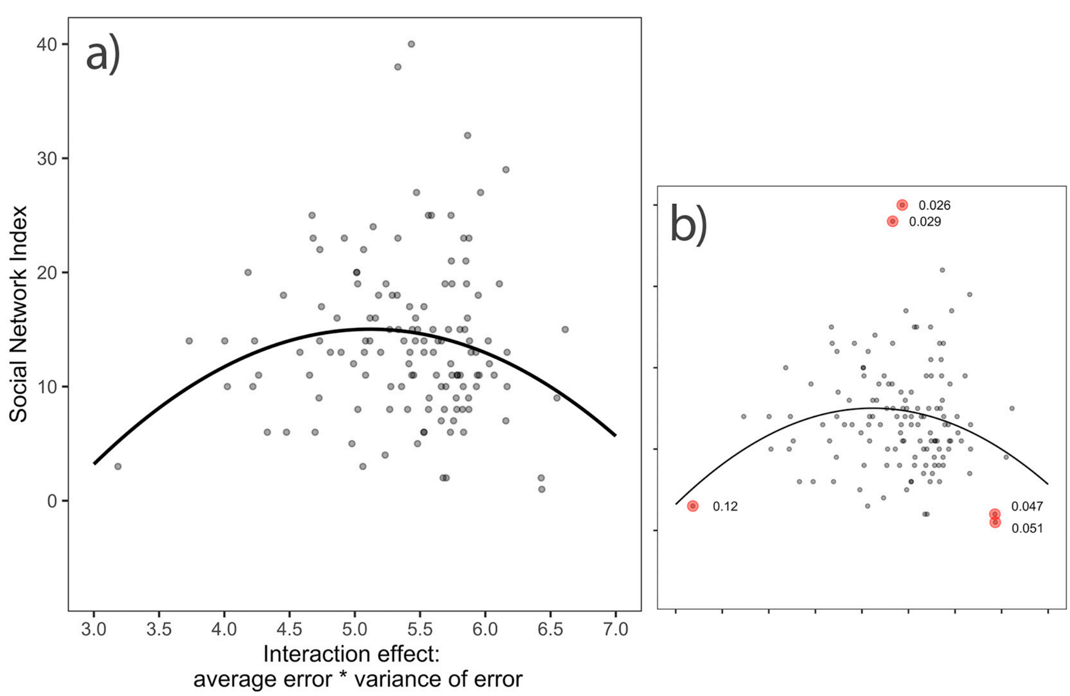

Fig 2. (a) The interaction between average error, and variance of error, as predicted by our three models. Each datapoint is a rater. A greater x-value signifies more error in their rating of stress, with a larger variance in error across their ratings. A smaller x-value signifies less error in their rating of stress, with less variance (or more consistence) across their ratings. Values along the y axis were log transformed to reduce skew in the data. (b) We identified several potential influential datapoints in our model ($n = 5$, circled in red). To assess how these are influencing our findings, we rebuilt our model with these individuals omitted. Values next to each of these datapoints represent the p value of the interaction between average error and variance of error, after omitting that individual. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in their ratings and a lower variance in their error (i.e., were consistently accurate) also reported a smaller social network. It was the more intermediately skilled individuals (i.e. with average accuracy) who reported the highest social network in our data.

However, we wish to highlight caution around these findings and present these analysis as exploratory. Visual inspection of the data suggests there may be some influential (or, outlying) data points (Fig. 2b, $n = 5$). We reanalysed the data with each of these individuals removed and in 2 out of 5 occasions, the interaction effect between average error and variance of error disappeared ($p = 0.120$ and $p = 0.051$); the effect held for the remaining 3. Therefore, we conclude that our model is not highly robust, and we should remain cautious about any interpretations.

Our four models assessing the effect of average error and variance of error on each of our expressivity measures extracted (overall expressivity, negative facet, positive facet and impulse strength), showed no significant relationships (all $p > 0.05$). Therefore, the ability to accurately detect stress in others was not associated with self-reported emotional expressivity. These models are fully described in the supplementary materials (SM3).

4. Discussion

Here, we found that those people who reported being more stressed were perceived as more stressed by others, and identified two different aspects of behaviour which appeared to influence this. First, those participants who showed a high proportion of displacement behaviours (i.e., where displacement behaviour accounted for a larger percentage of their total behaviour) were thought to be more stressed by raters. Second, those participants who showed a higher proportion of submissive behaviours were thought to be less stressed by raters. Total behaviours on the other hand, did not influence this rating, suggesting that it is not just the overall presence of any behaviour which leads to a perception of stress, but that there are some specific behavioural markers that raters were using informatively. Facial expressivity did not appear to influence stress ratings, however, an increase in both displacement activities and facial expressions increased how likable that individual was rated. When assessing if and how the stress-reading skills of our raters was associated with their social network size, our data suggests that our intermediately skilled raters (those who were not too accurate or inaccurate) reported having more social connections. This highlights a potential link between individual differences in signal processing and social network size. We found no associations with the raters' stress-reading skills and their own reported emotional expressivity.

Although a link between displacement behaviour and stress has been well documented, even within the context of the Trier Social Stress Test (Mohiyeddini & Semple, 2013), this is the first evidence that these behaviours are actually communicating stress to others in humans. Exactly what information is being communicated, and what adaptive value this information is still unknown. Stress is a psychological and physiological experience (Buchanan, Bagley, Stansfield, & Preston, 2012), but conceptualising stress in an emotional framework (i.e., considering the internal state of the actor) in the context of communication may not be helpful to fully understand why a signal may have emerged (Waller, Whitehouse, & Micheletta, 2017). Stress has a significant impact on behaviour and cognition, both increasing the propensity of certain behaviours (such as risk-taking, Porcelli & Delgado, 2017) in addition to impairing normal decision making and cognitive skills (Girotti et al., 2018). Therefore, it could be by recognising stress in others, this allows for an improved anticipation of an individual's future behaviour or the potential for that individual to behave unpredictably (in accordance with a behavioural ecology view, Crivelli & Fridlund, 2018). This could be highly adaptive for both the signaller and receiver if a product of this is a more coordinated social interaction with less uncertainty (Schmidt & Cohn, 2001). In non-human animals, we observe a link between the establishment of social relationships and stress; for example, when monkeys produce signals reinforcing a dominance hierarchy (e.g. produce a submissive behaviour towards a more dominant individual), both animals appeared less stressed as a result, as the other's motivation become less uncertain (Schino, Maestripieri, Turillazzi, & Scucchi, 1990). This is in some part comparable to the pattern observed in our data. Those individuals who produced more submissive-like behaviours were also perceived as less stressed. It could be that those individuals producing submissive-like behaviours appeared to have a more established relationship with the interviewer (or, more specifically, acknowledged the interviewer as the dominant individual in the interaction) and therefore appeared to have less uncertain and stressful social interaction.

Stressed individuals were rated as more likable. Displacements behaviours have been discussed quite extensively in behavioural contagion research, and both scratching and yawning have been reported to elicit a contagious effect in observers (Massen & Gallup, 2017; Platek, Mohamed, & Gallup, 2005; Schut, Grossman, Gieler, Kupfer, & Yosipovitch, 2015). One idea, although it has attracted some criticism (Massen & Gallup, 2017), is that this contagion has a social function by promoting an empathetic or pro-social response in others. If the individuals are inducing an empathetic-like response in the observers/raters, they may appear more likeable because of this. Or, it could be that an honest signal of weakness may represent an example of benign intent (Seyfarth & Cheney, 2017) and/or a willingness to engage in a cooperative rather than competitive interaction, something which could be a 'likable' or preferred trait in a social partner. Regardless of the mechanisms linking stress behaviour with likability, this finding represents a potential adaptive benefit to communicating stress. More likeable people may have more opportunity to develop social connections with others, build and maintain better social networks, and develop more friendships – something which has been demonstrated to have enormous fitness benefits in both human and non-human animals (Dunbar, 1988). Here, we also found that producing more facial expressions, behaviours which perhaps have a clearer signalling function, similarly increased how likable our actors were perceived. This fits with some of our current understanding of expressivity, which tends to suggest that people who are more "emotionally expressive" are more well-liked by others and have more positive social interactions (Riggio, 2017). We could therefore interpret our data as showing that stress behaviours are functioning in a similar way to that of facial expressions, and that they have the adaptive benefit of helping towards the development and maintenance of an individual's social network.

These findings feed well into the classic ideas presented by Tooby and Cosmides (1996), who considered the function of friendship within 'the banker's paradox'. In this framework, it is thought that we should seek out friendships that are *good investments:*, where the benefits received from the friendship outweigh the cost of the resources and effort spent to maintain it. As a consequence, behaviours which signal our trustworthiness and reliability should therefore be selected as adaptations during evolution, as well as those which honestly communicate the signallers' needs (as this makes it easier for others to invest efficiently, Delton & Robertson, 2012). The emergence of stress behaviour sits well within this model, if such behaviours are honestly communicating the need for assistance from others. Such a framework could also help to explain the interplay between stress behaviour, submissive behaviour, and rated stress. Those appearing more stressed in this study also behaved less submissively (i.e., produced less submission-associated behaviours). It is possible that those less submissive individuals (or, those individuals who are more dominant) may be expected to have a higher social status, and thus, they may represent a more valuable social connection. Therefore, affiliative or prosocial behaviours towards such individuals could be 'good investments' as the payoff of securing a social connection with this individual is greater. A proxy for how affiliative an individual will be towards another, or their willingness to provide another social support, could be the extent they report liking them. Thus, the positive relationship between stress behaviour and likability that we observe could be being further being modulated by this desire to support higher-status individuals (Tooby & Cosmides, 1996). In the bargaining model of depression presented by Hagen and Rosenström (2016), signalling distress is framed as a way that individuals can bargain for better treatment from social partners when faced with a reduction in social support. This strategy has been suggested as an adaptive function of depression; to appear vulnerable and be distant from others (i.e., withholding the social benefits they provide) which in turn can elicit a positive action in social partners. As the experience of depression co-occurs with the experience of stress/ anxiety, behaviours which function to communication these internal states may fit among this bargaining model of depression and help explain their presence in human behavioural repertoire.

For a signal to emerge, we should not only observe benefits for the signalling individual, but also benefits for the individual receiving and potentially responding to it (Laidre & Johnstone, 2013). Our data provided a chance to address this idea in an exploratory way. Here, we looked at how the stress-processing skills of our raters (i.e. how accurately they could detect stress in others) impacted their own social environment, with an expectation that those more skilled at reading others would report having more social connections. What we found however, was not a simple linear relationship such as this. Both the most and least accurate raters reported lower social networks, and it was in fact the more intermediately skilled (or, averagely skilled) raters which reported the most social connections. Those raters who were less accurate at reading stress (and thus may be communicatively less efficient) may find it more difficult to maintain and establish social relationships if they cannot communicate well with others (Riggio, 2017) and have a smaller social network as a consequence. Why those who were especially good at reading stress also reported smaller networks is less clear. First, there are many factors that contribute to social complexity aside from the size of a social network (Freeberg et al., 2012; Kappeler, 2019). Perhaps smaller networks are made up of higher quality social connections, and that this would also incur fitness benefits akin to simply having a larger network. If this was true, it could mean there are multiple strategies to maximise fitness (quality vs quantity of social partners), and something which could begin to explain why we can observe such individual variation in communicative ability (Kanai & Rees, 2011). Second, because not all communication is honest and in some cases can be used to manipulate others (Bradbury & Vehrencamp, 1998), being too accurate at reading the motivations of others may not be a desirable characteristic in a social partner. Therefore, this could compromise someone's ability to develop social connections, if their social partner can detect that their emotions or motivations are being detected when they prefer to conceal them. We do highlight some caution around these findings, however. Inspections of our models did suggest that they are not highly robust, so instead we wish to present these as results as ideas for future exploration as opposed to concrete findings.

In sum, we present evidence to suggest that people can quite effectively detect stress in others from behaviour alone, not just as present/ non present, but along a continuum. We hope these findings shed some light on the communicative functionality of stress behaviour and help explain why these behaviours have evolved. As we made many behavioural comparisons in this study beyond our main hypotheses, we wish to remind readers of the exploratory nature of this study, and we hope these findings may provide opportunities for more hypothesis-driven research in the future. Finally, we highlight a need for more research into individual differences in stress behaviour production and the social impact they have on both the actors and the receivers in order to further understand why these behaviours have evolved.

Competing interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgements

We would like to thank Holly Radford and Mariam Bahrainwala for collecting the videos used in this study. We would also like to thank our funders for supporting this research. This study was funded in part by a British Academy small research grant SRG18R1\180883. This study has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 865694 awarded to BW).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.evolhumbehav.2022.04.001.

References

Bender, R., & Lange, S. (2001). Adjusting for multiple testing—When and how? *Journal of Clinical Epidemiology, 54*(4), 343–349. https://doi.org/10.1016/S0895-4356(00)00314-0

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication* (Second). Sinauer Associates.

Bradshaw, R. H. (1993). Displacement activities as potential covert signals in primates. *Folia Primatologica; International Journal of Primatology, 61*(3), 174–176. https://doi.org/10.1159/000156746

Buchanan, T. W., Bagley, S. L., Stansfield, R. B., & Preston, S. D. (2012). The empathic, physiological resonance of stress. *Social Neuroscience, 7*(2), 191–201. https://doi.org/10.1080/17470919.2011.588723

Chance, M. R. A. (1962). An interpretation of some agonistic postures; the role of cut-off acts and postures in evolutionary aspects of animal communications, 1961. *Symposia Zool Soc London, 8*, 71–89.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M. (1997). Social ties and susceptibility to the common cold. *Journal of the American Medical Association, 277*, 1940–1944.

Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences, 22*(5), 388–399. https://doi.org/10.1016/j.tics.2018.02.006

Delton, A. W., & Robertson, T. F. (2012). The social cognition of social foraging: Partner selection by underlying valuation. *Evolution and Human Behavior: Official Journal of the Human Behavior and Evolution Society, 33*(6), 715–725. https://doi.org/10.1016/j.evolhumbehav.2012.05.007

Dezecache, G., Jacob, P., & Grèzes, J. (2015). Emotional contagion: Its scope and limits. *Trends in Cognitive Sciences, 19*(6), 297–299. https://doi.org/10.1016/j.tics.2015.03.011

Duesenberg, M., Weber, J., Schulze, L., Schaeuffele, C., Roepke, S., Hellmann-Regen, J., Otte, C., & Wingenfeld, K. (2016). Does cortisol modulate emotion recognition and empathy? *Psychoneuroendocrinology, 66*, 221–227. https://doi.org/10.1016/j.psyneuen.2016.01.011

Dunbar, R. I. M. (1988). *Primate social systems*. Springer Science & Business Media.

Ekman, P., & Friesen, W. (1978). *Facial Action Coding System* (Vol. 1). Consultion Psychologists Press.

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 367*(1597), 1785–1801. https://doi.org/10.1098/rstb.2011.0213

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution, 7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Giannakakis, G., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P. G., … Tsiknakis, M. (2017). Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control, 31*, 89–101. https://doi.org/10.1016/j.bspc.2016.06.020

Girotti, M., Adler, S. M., Bulin, S. E., Fucich, E. A., Paredes, D., & Morilak, D. A. (2018). Prefrontal cortex executive processes affected by stress in health and disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 85*, 161–179. https://doi.org/10.1016/j.pnpbp.2017.07.004

Gross, J. J., & John, O. E. (1997). Revealing feelings: Facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of Personality and Social Psychology, 72*(2), 435–448.

Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour, 42*(1), 1–14. https://doi.org/10.1016/S0003-3472(05)80600-1

Hagen, E., & Rosenström, T. (2016). Explaining the sex difference in depression with a unified bargaining model of anger and depression. *Evolution, Medicine, and Public Health*. https://doi.org/10.1093/emph/eow006

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience, 12*(4), 231–242. https://doi.org/10.1038/nrn3000

Kappeler, P. M. (2019). A framework for studying social complexity. *Behavioral Ecology and Sociobiology, 73*(1), 13. https://doi.org/10.1007/s00265-018-2601-8

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' – A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1–2), 76–81. https://doi.org/10.1159/000119004

Laidre, M. E., & Johnstone, R. A. (2013). Animal signals. *Current Biology, 23*(18), R829–R833. https://doi.org/10.1016/j.cub.2013.07.070

Maestripieri, D., Schino, G., Aureli, F., & Troisi, A. (1992). A modest proposal: Displacement activities as an indicator of emotions in primates. *Animal Behaviour, 44* (5), 967–979. https://doi.org/10.1016/S0003-3472(05)80592-5

Massen, J. J. M., & Gallup, A. C. (2017). Why contagious yawning does not (yet) equate to empathy. *Neuroscience & Biobehavioral Reviews, 80*, 573–585. https://doi.org/10.1016/j.neubiorev.2017.07.006

Mayo, L. M., & Heilig, M. (2019). In the face of stress: Interpreting individual differences in stress-induced facial expressions. *Neurobiology of Stress, 10*, 100166. https://doi.org/10.1016/j.ynstr.2019.100166

McFarland, D. J. (1966). On the causal and functional significance of displacement activities. *Zeitschrift für Tierpsychologie, 23*(2), 217–235. https://doi.org/10.1111/j.1439-0310.1966.tb01600.x

Mohiyeddini, C., Bauer, S., & Semple, S. (2013). Displacement behaviour is associated with reduced stress levels among men but not women. *PLoS One, 8*(2), Article e56355. https://doi.org/10.1371/journal.pone.0056355

Mohiyeddini, C., Bauer, S., & Semple, S. (2015). Neuroticism and stress: The role of displacement behavior. *Anxiety, Stress, and Coping, 28*(4), 391–407. https://doi.org/10.1080/10615806.2014.1000878

Mohiyeddini, C., & Semple, S. (2013). Displacement behaviour regulates the experience of stress in men. *Stress, 16*(2), 163–171. https://doi.org/10.3109/10253890.2012.707709

Morgan, D. L. (1988). Age differences in social network participation. *Journal of Gerontology, 43*(4), S129–S137. https://doi.org/10.1093/geronj/43.4.S129

Pico-Alfonso, M. A., Mastorci, F., Ceresini, G., Ceda, G. P., Manghi, M., Pino, O., … Sgoifo, A. (2007). Acute psychosocial challenge and cardiac autonomic response in women: The role of estrogens, corticosteroids, and behavioral coping styles. *Psychoneuroendocrinology, 32*(5), 451–463. https://doi.org/10.1016/j.psyneuen.2007.02.009

Platek, S. M., Mohamed, F. B., & Gallup, G. G. (2005). Contagious yawning and the brain. *Cognitive Brain Research, 23*(2–3), 448–452. https://doi.org/10.1016/j.cogbrainres.2004.11.011

Porcelli, A. J., & Delgado, M. R. (2017). Stress and decision making: Effects on valuation, learning, and risk-taking. *Current Opinion in Behavioral Sciences, 14*, 33–39. https://doi.org/10.1016/j.cobeha.2016.11.015

Preston, S. D., & de Waal, F. B. M. (2002). The communication of emotions and the possibility of empathy in animals. In *Altruism & altruistic love: Science, philosophy, & religion in dialogue* (pp. 284–308). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195143584.003.0025.

Riggio, H. R. (2017). Emotional Expressiveness. In V. Zeigler-Hill, & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1–8). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_508-1.

Scheider, L., Liebal, K., Oña, L., Burrows, A., & Waller, B. (2014). A comparison of facial expression properties in five hylobatid species: Facial expressions in hylobatids. *American Journal of Primatology, 76*(7), 618–628. https://doi.org/10.1002/ajp.22255

Schino, G., Maestripieri, D., Turillazzi, P. G., & Scucchi, S. (1990). Social tension in familiar and unfamiliar pairs of long-tailed macaques. *Behaviour, 113*(3–4), 264–272. https://doi.org/10.1163/156853990X00518

Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *American Journal of Physical Anthropology, 116*(S33), 3–24. https://doi.org/10.1002/ajpa.20001

Schut, C., Grossman, S., Gieler, U., Kupfer, J., & Yosipovitch, G. (2015). Contagious itch: What we know and what we would like to know. *Frontiers in Human Neuroscience, 9*. https://doi.org/10.3389/fnhum.2015.00057

Scott-Phillips, T. C., Blythe, R. A., Gardner, A., & West, S. A. (2012). How do communication systems emerge? *Proceedings of the Royal Society B: Biological Sciences, 279*(1735), 1943–1949. https://doi.org/10.1098/rspb.2011.2181

Seyfarth, R. M., & Cheney, D. L. (2017). The origin of meaning in animal signals. *Animal Behaviour, 124*, 339–346. https://doi.org/10.1016/j.anbehav.2016.05.020

Spielberger, C. D. (1983). *Manual for the state-trait anxiety inventory: STAI (Form Y)*. Consulting Psychologists Press.

Tinbergen, N. (1952). 'Derived' activities; their causation, biological significance, origin, and emancipation during evolution. *The Quarterly Review of Biology, 27*(1), 1–32.

Tomasello, M. (2010). *Origins of human communication*. MIT press.

Tooby, J., & Cosmides, L. (1996). Friendship and the Banker's Paradox: Other pathways to the evolution of adaptations for altruism. In *, Vol. 88. Evolution of social behaviour patterns in primates and man* (pp. 119–143). Proceedings of the British Academy.

Troisi, A. (1999). Ethological research in clinical psychiatry: The study of nonverbal behavior during interviews. *Neuroscience & Biobehavioral Reviews, 23*(7), 905–913. https://doi.org/10.1016/S0149-7634(99)00024-X

Troisi, A. (2002). Displacement activities as a behavioral measure of stress in nonhuman primates and human subjects. *Stress, 5*(1), 47–54. https://doi.org/10.1080/102538902900012378

Waller, B. M., Whitehouse, J., & Micheletta, J. (2017). Rethinking primate facial expression: A predictive framework. *Neuroscience & Biobehavioral Reviews, 82*, 13–21. https://doi.org/10.1016/j.neubiorev.2016.09.005

Whitehouse, J., Micheletta, J., & Waller, B. M. (2017). Stress behaviours buffer macaques from aggression. *Scientific Reports, 7*(1), 11083. https://doi.org/10.1038/s41598-017-10754-8