Manuscript accepted for publication in *Cortex*, 12.10.16.

The neuropsychology of first impressions: Evidence from Huntington's disease

Reiner Sprengelmeyer (1,4*), Andrew W. Young (2), Eva-Maria Baldas (1), Iris Ratheiser (1), Clare A.M. Sutherland (2,5), Hans-Peter Müller (1), Georg Grön (3), Sigurd D. Süssmuth (1), G. Bernhard Landwehrmeyer (1), Michael Orth (1)

(1) Department of Neurology, University of Ulm, Ulm, Germany
(2) Department of Psychology, University of York, York, England
(3) Department of Psychiatry, University of Ulm, Ulm, Germany
(4) School of Psychology and Neuroscience, University of St Andrews, St Andrews, Scotland
(5) School of Psychology, University of Western Austraila, Australia

Corresponding author:
R Sprengelmeyer
School of Psychology and Neuroscience
University of St Andrews
St Andrews
Scotland

Email: rhs3@st-and.ac.uk

ABSTRACT

Impairments of emotion recognition have been widely documented in Huntington's disease (HD), but little is known concerning how these relate to other aspects of social cognition, including first impressions of traits such as trustworthiness and dominance. Here, we introduce a novel and sensitive method to investigate the ability to evaluate trustworthiness and dominance from facial appearance, with control tasks measuring ability to perceive differences between comparable stimuli. We used this new method together with standard tests of face perception to investigate social cognition in HD. We found that a subgroup of people with HD was impaired at perceiving trustworthiness and dominance, and that perceiving trustworthiness and dominance were correlated with impaired facial expression recognition. In addition, we used diffusion tensor imaging (DTI) to provisionally identify candidate brain regions associated with social cognition by contrasting regional functional anisotropy (FA) measures between subgroups of HD participants showing normal or impaired perception of trustworthiness and dominance, and by correlating these regional brain abnormalities with behavioural performance on tests of emotion recognition. In this way we show for the first time alterations in perception of trustworthiness and dominance in people with HD and link these to regions which may map the boundaries of the social brain. The pattern of breakdown seen in this neurodegenerative disease can thus be used to explore potential inter-relationships between different components of social cognition.

Keywords: Huntington's disease, trustworthiness, dominance, face perception, emotion recognition, DTI

1 INTRODUCTION

Huntington's disease (HD) is a hereditary, progressive neurodegenerative disorder, caused by a single gene mutation on chromosome 4. The neuropathology of HD is widespread and variable, affecting predominantly the basal ganglia, but other cortical regions as well, albeit to a lesser extent. The disorder is characterised by involuntary choreiform movements, cognitive deterioration (Dumas, van den Bogaard, Middelkoop, & Roos, 2013; Papoutsi, Labuschagne, Tabrizi, & Stout, 2014; Paulsen, 2011), affective disturbances (Duff, Paulsen, Beglinger, Langbehn, & Stout, 2007; Sprengelmeyer, et al., 2014), and impaired emotion processing (Rees, et al., 2014; Sprengelmeyer, Schroeder, Young, & Epplen, 2006; Sprengelmeyer, et al., 1996).

Facial emotion recognition has been the most widely studied aspect of social cognition in HD. A number of studies have found a disproportionately severe deficit in disgust recognition in the pre-symptomatic (Gray, Young, Barker, Curtis, & Gibson, 1997; Hennenlotter, et al., 2004; Sprengelmeyer, et al., 2006) and symptomatic stages of the disorder (Rees, et al., 2014; Sprengelmeyer, et al., 1996; Wang, Hoosain, Yang, Meng, & Wang, 2003), and there are findings indicating that the experience of disgust itself can also be affected (Hayes, Stevenson, & Coltheart, 2007; Mitchell, Heims, Neville, & Rickards, 2005). In these studies, however, the disproportionately severe deficit with disgust is mostly found within the context of more widespread problems in emotion recognition, and especially recognition of anger (Calder, et al., 2010). Some studies however only find a more generalized deficit, with all negative emotions affected (Johnson, et al., 2007; Milders, Crawford, Lamb, & Simpson, 2003).

Although studies of emotion recognition in HD have thus generated mixed findings concerning *which* emotions are most severely affected, all studies find some degree of overall impairment in facial expression recognition. However, social cognition entails more than recognising facial expressions. In particular, in social encounters people make rapid inferences about strangers derived from a multiplicity of cues from their faces, voices, body shapes, posture and gestures. These first impressions cover a wide range of traits that have been most extensively studied in the case of first impressions derived from facial cues. It is now well-established that these impressions are easily created by faces and that they can affect real-life decisions (Castle, et al., 2012; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015), but nothing is at present known about how HD affects the ability to form these social trait impressions. This is an important issue both clinically and scientifically. Its clinical importance stems from the fact that people with HD need to rely increasingly on others the further their disease progresses and the more they lose their autonomy. Failure to appropriately interpret social signals and form impressions of others may therefore negatively affect both people with HD and their carers. Scientifically, the importance lies in the possibility of using the pattern of breakdown seen in a neurodegenerative disease to investigate the inter-relationships between different abilities involved in social cognition. This tactic of investigating patterns of breakdown in neurodegenerative disorders has been very fruitful in other areas (Hoffman, Meteyard, & Patterson, 2014; Patterson & Hodges, 1992).

The scientific part of this agenda depends heavily on identifying the most appropriate theoretical questions. An important advance in understanding how neurologically normal individuals derive first impressions from facial cues has come from the work of Oosterhof and Todorov (Oosterhof & Todorov, 2008), who used Principal Component Analysis to find the underlying structure of 15 traits commonly perceived in faces. Their analysis showed that these traits could be organised into a two-dimensional space in which the first principal component closely corresponded to perceived trustworthiness while the second principal component was approximated by perceived dominance.

Although Oosterhof and Todorov's (2008) work clearly establishes the usefulness of a dimensional approach to first impressions derived from facial cues, it does not in itself establish where these dimensions come from. An appealing hypothesis is that they relate to fundamental mechanisms of social threat appraisal present in other primate species (Fiske, Cuddy, & Glick, 2007; Olivola, Funk, & Todorov, 2014), whereby conspecifics are evaluated in terms of their likely intention to help or harm (trustworthiness) and their ability to carry out such intentions (dominance). Inferences about trustworthiness and dominance from faces can therefore result in important behavioural responses such as approaching or avoiding a person (Winston, Strange, O'Doherty, & Dolan, 2002).

Against this background, the relation between social trait inferences and facial expression recognition is an interesting question. Although research has often looked separately at facial expression recognition and social inferences from faces, a growing body of evidence points to a possible relationship between both processes, as summarised in the emotion overgeneralisation hypothesis (Montepare & Dobish, 2003). This behaviourally established relationship suggests that facial expression recognition and inferences of trustworthiness may share some of the same neural mechanisms (Mattavelli, Andrews, Asghar, Towler, & Young, 2012; Said, Haxby, & Todorov, 2011). Importantly, trustworthiness ratings of faces by neurologically normal perceivers are particularly low when the faces show expressions of disgust and anger (Willis, Palermo, & Burke, 2011; Sutherland, Young, & Rhodes, in press). These are emotions whose recognition, as noted above, is often particularly affected in HD.

Oosterhof & Todorov's (2008) perspective offers insight into why there might be common neural mechanisms for perceiving facial expressions of emotion and trustworthiness. In particular, the good or bad intentions inferred in evaluating the trustworthiness dimension may be linked either to an overt emotional expression (Secord, 1958; Sutherland, et al., 2013) or to the subtle resemblance of a target's static physiognomy to an emotional expression (Oosterhof & Todorov, 2008). For example, an overt expression of happiness or the subtle resemblance of a neutral face to a happy expression is judged as trustworthy, while an overt or subtle resemblance to an angry expression is judged as untrustworthy (Adams, Nelson, Soto, Hess, & Kleck, 2012; Montepare & Dobish, 2003; Oosterhof & Todorov, 2008; Zebrowitz, Kikuchi, & Fellous, 2007).

At present, the relation of the dominance dimension to facial expressions of basic emotion is less clear. In Oosterhof and Todorov's (2008) model, this dimension is suggested to represent a judgement of a target's capability of carrying out their intentions based on cues to physical strength, masculinity and maturity. However, this idea does not rule out a possible influence of facial expression alongside these more structural facial characteristics.

Indeed support for a link to facial emotion recognition comes from studies showing correlations between facial emotion recognition and trustworthiness and dominance inferences made from neutral faces (Montepare & Dobish, 2003, 2014; Oosterhof & Todorov, 2008). If the emotion overgeneralisation hypothesis holds true, we would therefore expect people with deficits in emotion recognition (such as those noted in HD) to be impaired in evaluating both trustworthiness and dominance from the faces of social counterparts.

The current study therefore investigates for the first time the ability to evaluate trustworthiness and dominance in people with HD. To achieve a sensitive behavioural measure well-suited to this purpose, we used computer image manipulation methods equivalent to those validated in previous studies (Mattavelli, et al., 2012; Sutherland, et al., 2013) to create continua of images that varied in perceived trustworthiness or dominance. We then measured how well these subtle gradations in manipulated trait impressions were perceived by people with HD and by control participants, and established the relation between performance deficits and other behavioural measures including measures of emotion recognition.

Using the data from this novel, sensitive behavioural measure we were able to identify subgroups of HD participants with relatively intact or impaired perception of trustworthiness and dominance. We then explored differences between these subgroups in terms of other behavioural tests (including ability to recognise facially expressed emotion) and with regional functional anisotropy (FA) measures based on diffusion tensor imaging (DTI).

2 METHODS

2.1 Participants

2.1.1 HD participants and controls

We tested 67 participants (41 HD participants, 26 controls). All HD participants were recruited from the Department of Neurology, University of Ulm, Germany. They underwent routine motor, neuropsychiatric and cognitive testing and a detailed assessment of various aspects of face perception and social cognition (described below). Fifty-eight of the 67 participants (32 HD participants, 26 controls) also underwent 3T MRI T1 and diffusion-tensor (DTI) imaging. The project was approved by the local ethics committee and written informed consent was obtained from all participants according to ICH-Guidelines.

Eligible participants with manifest HD were at least 18 years old with a CAG-repeat length of 40 or more. All participants with HD had unequivocal mild to moderate motor signs. Exclusion criteria were current psychiatric disorders (such as Major Depressive Disorder, psychosis, or drug and alcohol abuse as defined by DSM-IV-TR criteria), past and current neurological disorders other than HD, head injury, current participation in a clinical drug trial, and not being able to tolerate or safely undergo MRI. The same exclusion criteria applied to controls. Control participants were partners and friends of HD participants or were recruited from hospital staff.

2.1.2 Clinical assessment

Motor signs, the ability to cope with the demands of daily life, affect, and cognition were assessed with the Unified Huntington's Disease Rating Scale - UHDRS: (Huntington-Study-Group, 1996).

The UHDRS Motor score (range between 0 = unimpaired, and 124 = severely impaired) summarises clinical ratings of voluntary and involuntary movements. The UHDRS Total Functioning Capacity score (TFC: range between 13 = unimpaired, and 0 = severely impaired) reflects the ability of participants to handle occupational, domestic, and financial matters.

2.1.3 Participant charateristics

Background demographic and clinical details of the participant groups are summarised in Table 1a.

Table 1a about here, please

There was no significant difference between HD participants and controls with respect to age (*t*(65) = 0.71, *p* = .48, *d* = 0.17), years of education (*t*(65) = -1.29, *p=.20 d* = 0.31), and IQ as measured with the MWT-B (Lehrl, 2005) (*t*(65) = -1.46, *p* = .15, *d* = 0.36). In a refined data analysis, we compared the performance and DTI imaging of HD participants who were impaired on the trust and dominance recognition task with HD participants who showed no impairment in trust and dominance recognition. Clinical details for the HD subgroups can be found in Table 1b, and under the heading "Subgrouping of HD participants" in the results section.

Face perception and social cognition tasks

We created a new measure of how participants evaluated trustworthiness and dominance from facial appearance, with control tasks measuring ability to perceive differences in apparent gender between comparable stimuli. To complement these new measures we used standard tests of ability to perceive unfamiliar faces (Benton test), recognise subtle moods and feelings (Mind in the Eyes test), and recognise basic emotions (Ekman 60 Faces test). These are described below.

Trustworthiness, dominance and gender perception task: This new task was created for this project, so we describe it in detail. Our aim was to devise a task that could measure the evaluation of social traits of trustworthiness and dominance that correspond to the axes of Oosterhof and Todorov's (2008) model, and to include matched conditions involving the evaluation of face gender as a point of comparison based on purely physical characteristics.

Decisions about a person's trustworthiness or dominance based on their face are of course of low validity, though many researchers take seriously the possibility that there may be an underlying 'kernel of truth' (Bruce & Young, 2012). Importantly, however, regardless to what extent such judgements are valid there is a remarkably good degree of agreement between neurologically normal observers concerning which faces *look* trustworthy or *look* dominant. We therefore chose to measure the perception of trustworthiness and dominance in HD in terms of the extent to which the judgements of participants with HD correlated with those of neurologically normal participants.

Because the facial cues that underlie perception of trustworthiness and dominance are likely multiple in nature and not yet fully understood, we used an approach based on averaging everyday images of male or female faces to create prototypes that look trustworthy or untrustworthy, dominant or non-dominant, and then morphing and caricaturing between these prototypes to create quasi-linear continua varying on trustworthiness, dominance, or gender. This approach creates images that capture a range of everyday cues to the appropriate trait, and has been validated in previous studies (Mattavelli, et al., 2012; Oldmeadow, Sutherland, & Young, 2013; Sutherland et al., 2013; Sutherland, Young, Mootz, & Oldmeadow, 2015; Vernon, Sutherland, Young, & Hartley, 2014).

Figure 1a and 1b about here, please

We begin by describing the stimuli created for the trustworthiness task and its matched gender perception task. These tasks were based on the matrix of face-like images shown in Figure 1a, in which the images vary across 10 levels of perceived untrustworthiness to trustworthiness along the horizontal dimension and 10 levels of perceived masculinity to femininity in the vertical dimension.

The matrix was created as follows, using similar procedures to a previous study (Mattavelli, et al., 2012). Photographs of 500 adult male and 500 adult female faces were collected from the internet. The photographs varied in pose, age and expression, to allow a wide a range of cues to be present in the images. However, photographs of famous people were excluded, to eliminate potential influences of prior knowledge about the person. Moreover, only Caucasian adult faces were chosen, to eliminate potential cross-cultural confounds. From ratings of the perceived trustworthiness of these 1,000 face photographs established in previous studies (Santos & Young, 2005; Santos & Young, 2008; Santos & Young, 2011; Sutherland, et al., 2013; Sutherland, et al., 2015) the 15 male faces with lowest rated trustworthiness and the 15 male faces with highest rated trustworthiness were selected, subject to constraints that the photographs included no spectacles, were sufficiently close to frontal view that both eyes were visible, showed no beards or moustaches, and that there were no more than two faces with hats in each set. These constraints were introduced only to allow the creation of relatively sharp averaged images. There was no matching on any other characteristics, with free variation of all other aspects.

The faces in each set of 15 photographs were then averaged using PsychoMorph software (Tiddeman, Perrett, & Burt, 2001) to generate averaged images of low trustworthiness and high trustworthiness male faces, and a graded image continuum between these extremes was created by image morphing (see (Sutherland, et al., 2013), Figure 1, for images created by an equivalent procedure). In exactly the same way, we also created a continuum of images between averages of the 15 female faces with lowest rated trustworthiness and the 15 female faces with highest rated trustworthiness. These procedures therefore resulted in sets of male and female face-like images of varying apparent trustworthiness, and the trustworthiness of these was rated by a group of neurologically normal perceivers. From these ratings, we selected a male and a female face-like image with matched low rated trustworthiness, and a male and female image with matched high rated trustworthiness. These images formed the prototypes used to generate the trustworthiness by gender image matrix, and are shown marked by red frames in Figure 1a.

Image continua were then created for trustworthiness of male and female faces (from very low to very high trustworthiness) by slightly caricaturing the prototype at each level of trustworthiness to increase its distance from the other prototype with the same gender, and by morphing between the corresponding high and low prototypes in seven steps. In this way, a quasi-linear continuum of 10 male face-like images of varying trustworthiness was created (forming the second row in Figure 1a), and a corresponding continuum of 10 female face-like images of varying trustworthiness (the ninth row in Figure 1a). By applying the same process to the 10 images from each of these horizontal continua from Figure 1a in the vertical dimension, the full 10x10 matrix of face-like images was created. Our intention was that as far as possible perceived trustworthiness should change primarily along the horizontal dimension in this matrix, and perceived gender should change primarily along the vertical. Pilot data with neurologically normal participants confirmed that this was the case.

To measure their perception of trustworthiness, participants were shown 40 of the stimuli from Figure 1a, one at a time in an unpredictable order. These 40 stimuli comprised the images from the second, fourth, seventh and ninth columns of Figure 1a. There were therefore four different stimuli (one from each of the four columns) at ten different levels of trustworthiness (corresponding to the rows of Figure 1a). Participants were asked the question "Ist diese Person vertrauenswürdig? - Is this person trustworthy?" and invited to rate each image on a 1-7 Likert scale from "1 – Nein! " (no!) to "7 – Sehr! " (very much!) (see Figure 2a) . Each face remained on the computer screen until the participant had responded verbally, and responses were noted down by the experimenter. Each block began with eight practice trials to familiarise participants with the task. These practice trials were excluded from the analysis.

To measure the perception of trustworthiness in this task, we averaged the ratings across the four images at the same level of trustworthiness, to yield a single set of 10 mean ratings; with one mean rating for each level of the 1-10 scale represented by the columns in Figure 1a. Each participant's mean ratings were then correlated with the 1-10 scale values. In this way a high correlation coefficient would imply that images with higher levels of computer-manipulated trustworthiness were rated as higher in perceived trustworthiness. These correlation coefficients were subjected to Fisher r to z transformations for analysis. Note that this method for evaluating how trustworthiness is perceived involves determining how well a participant's ratings fit the quasi-linear scaling of the computer-manipulated images. For this reason, we used the correlation as the measure of how well the linearity of the computer-manipulated continuum is perceived, not the ratings themselves. This use of a correlation as an estimate of the underlying linearity of responses eliminates problems that might otherwise arise from differences in the way in which participants use what is in essence a subjective rating scale (for example, one participant might use the full 1-7 range of ratings whilst another only uses the range 3-7, but if both show responses that are evenly graded across the continuum this will not affect the correlation).

An equivalent task was used to measure each participant's perception of femininity. For this, the 40 stimuli comprised the second, fourth, seventh and ninth rows in Figure 1a. Participants were asked "Erscheint diese Person weiblich? - Does this person look feminine?" and the 1-7 rating scale ranged from "1 – Nein! " (no!) to "7 – Sehr! " (very much!) (see Figure 2b). These perceived femininity ratings were now averaged across the four images at the same level of femininity, to yield a single set of 10 mean ratings; with one mean rating for each level of the 1-10 scale represented by the rows in Figure 1a. Each participant's mean ratings were then correlated with the 1-10 scale values, so that a high correlation coefficient would imply that images with increasingly high levels of computer-manipulated masculinity-femininity were rated as increasingly high in perceived femininity. These correlation coefficients were subjected to Fisher r to z transformations for analysis. The femininity rating task served as a control comparison condition to help evaluate whether impairments on the trustworthiness rating task reflect more general deficits in face perception, as it requires the evaluation of small physical differences between the images that are not related to trustworthiness *per se*.

To measure perceived dominance, comparable methods were used to create the matrix shown in Figure 1b, in which perceived dominance changes primarily along the horizontal dimension in this matrix, and perceived gender changes primarily along the vertical. As before, this was achieved by matching the rated dominance of computer-manipulated images that were created from averages of low rated dominance and high rated dominance male and female face photographs. The low and high dominance male and female prototypes are marked with red frames in Figure 1b.

From this matrix, we used the images from the second, fourth, seventh and ninth columns for a 1-7 dominance rating task using the question "Erscheint diese Person herrisch? – Does this person look dominant?" (see Figure 2c) and we again created a femininity rating comparison task with stimuli from the second, fourth, seventh and ninth rows. Analyses used the correlational method already described (see Figure 2d).

Figure 2a to 2d about here, please

Overall, then, participants completed four blocks of trials, each involving rating 40 images. These blocks were presented in a fixed order, with the first block involving trustworthiness ratings, the second block rating the degree of femininity of the trustworthiness matrix control images, the third block rating dominance, and the fourth block rating the degree of femininity of the dominance matrix control images**.** Our reasons for using a fixed order of presentation were that we were uncertain of the extent to which these novel tasks might prove fatiguing for our HD participants and we therefore prioritised the trustworthiness perception measure. Trustworthiness is a characteristic that perceivers find particularly salient (Todorov et al., 2015) and as noted in our Introduction it has already been linked to the analysis of expression.

Benton Facial Recognition Test: We used the short form of the Benton Facial Recognition Test (Benton, Sivan, Hamsher, Varney, & Spreen, 1994). Participants had to choose which of 6 photographs of unfamiliar faces are pictures of the same person as a simultaneously presented target face photograph. The short form includes 13 items involving choice of identical photographs, as well as transformations of orientation or lighting, which are pooled to give an overall total score.

Mind in the Eyes Test: This task (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) consists of 32 experimental trials and one practice trial. In each trial, the eye region of a male or female face was shown. Participants were asked to decide what the person in the picture is thinking or feeling by choosing one of four words that best describes the depicted person's state of mind. The four alternative descriptors were given at the top and bottom of the screen.

Ekman 60 Faces Test: Photographs of the faces of 10 people (6 female, 4 male) were selected from the (Ekman & Friesen, 1976) series; for details see the handbook of the FEEST – Facial Expressions of Emotion: Stimuli and Tests (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). For each face, there were poses corresponding to each of 6 emotions (happiness, surprise, fear, sadness, disgust, and anger), giving a total of 60 photographs. These were shown one at a time in pseudo-random order on a computer screen, each for an unlimited time until a response was made. On each trial, the participant had to decide which of the emotion names best described the facial expression shown. The names of the six emotions were given in German at the bottom of the screen (Freude - happiness, Erstaunen - surprise, Angst - fear, Trauer - sadness, Ekel - disgust, Wut – anger), and these words were available throughout the test. There were 6 practice trials and 60 test trials (one for each of the 6 emotions across the 10 posers), leading to an overall accuracy score out of 60 and a possible maximum of 10 for recognising each of the six emotions.

Magnetic resonance imaging

MRI data from 32 of the 41 HD participants involved in this study, and from all 26 controls were acquired using a 3 Tesla Siemens Magnetom Allegra. We used the DTI analysis software 'TIFT - Tensor Imaging and Fiber Tracking' (Mueller, et al., 2007) for image post-processing and for the statistical comparison of FA (functional anisotropy) maps. Prior to FA-calculation, motion artifacts were eliminated separately in each volume and each participant (H. P. Mueller, et al., 2011). All individual data sets were aligned to the AC-PC line and linearly transformed to fit with manually set landmarks. In order to spatially normalise the scans on the Montreal Neurological Institute (MNI) stereotactic standard space (Brett, Johnsrude, & Owen, 2002), both a (b=0)-template and an FA-template had to be created (Unrath, et al., 2010). The first (b=0)-template was created by arithmetically averaging the (b=0)-volumes of all participants. From averaging all FA maps of landmark-normalised data sets, the first FA template was calculated. Nonlinear MNI normalisation was performed using iteratively b=0 and FA templates. The resulting MNI normalised data sets were then used for creating improved b=0 and FA templates, obtained from the previous normalisation step. The whole process was iterative and stopped when the correlation of all individual MNI normalised b=0 images with the b=0 template reached r>.70.

After normalisation, all individual DTI data sets were used to calculate the second-rank diffusion tensor, the Eigenvalues, the Eigenvectors, and the FA for quantification of the diffusion anisotropy (Basser & Jones, 2002). For subsequent smoothing, we used an 8 mm full width at half maximum Gaussian filter (Unrath, et al., 2010). FA maps of HD groups were compared voxel-wise using t-tests. FA values below 0.2 were excluded from analysis (Kunimatsu, et al., 2004). Statistical results were corrected for multiple comparisons using the false-discovery-rate (FDR) algorithm (Genovese, Lazar, & Nichols, 2002) with a significance level set at p<.05. To further reduce the probability of a type 1 error, we eliminated clusters smaller than 268 voxels (which corresponds to a sphere with radius 4 mm).

RESULTS

Perception of trustworthiness and dominance

Our key aim was to use the newly developed measures of perception of trustworthiness and dominance to investigate the breakdown of social cognition in Huntington's disease. These tasks used computer image manipulation techniques to create quasi-linear continua of face-like images varying in perceived trustworthiness, dominance, or gender. We then measured how accurately the gradations in these continua were perceived by correlating each participant's ratings with the level of the manipulated trait. Overall performance was then compared between the HD and normal comparison groups based on the Fisher z-transformed correlations from the two conditions of the trustworthiness task (1. trustworthiness rating, 2. femininity rating], and the two conditions of the dominance task (1. dominance rating, 2. femininity rating).

Data were analysed with a mixed ANOVA, with task (i.e the use of stimuli and corresponding ratings based on the trustworthiness x gender vs the dominance x gender matrices) and condition (social ratings of trustworthiness or dominance vs the gender control ratings) as within-subject factors, and group (HD and controls) as a between group factor. This gave a significant effect of condition ($F(1,65) = 7.38$, $p = .008$, $\eta_p^2 = .102$), with better overall performance at rating gender compared to the more 'social' traits of trustworthiness and dominance, a significant effect of group ($F(1,65) = 10.9$, $p = .002$, $\eta_p^2 = .143$), with better overall performance by normal than HD participants, and a significant condition x group interaction ($F(1,65) = 4.75$, $p = .033$, $\eta_p^2 = .068$) which we decompose and interpret below. The effect of task ($F(1,65) = 0.824$, $p = .37$, $\eta_p^2 = .013$), the task x group interaction ($F(1,65) = 2.48$, $p = .12$, $\eta_p^2 = .037$), the task x condition interaction ($F(1,65) = 1.42$, $p = .24$, $\eta_p^2 = .021$), and the task x condition x group interaction ($F(1,65) = 0.13$, $p = .72$, $\eta_p^2 = .002$) were not significant. For details see Figure 3a and 3b.

The most important finding was the significant condition x group interaction. To explore this in more detail, we used t-tests to decompose the interaction. These revealed a significant group difference for the trustworthiness rating condition with better performance by normal than HD participants ($t(61.8) = 2.28$, $p = .026$, $d = 0.52$), no significant group difference for the trustworthiness control task (femininity rating) ($t(65) = .89$, $p = .38$, $d = 0.24$), a significant group difference for the dominance rating condition with better performance by normal than HD participants ($t(58.2) = 4.00$, $p < .001$, $d = 0.92$) and, owing to the small variances, a significant group difference for the dominance control task (femininity rating) with slightly better performance by normal than HD participants ($t(65) = 2.82$, $p = .006$, $d = 0.72$).

Figure 3a to 3d about here, please

Subgrouping of HD participants

As well as these statistically significant overall group differences, however, inspection of the raw data suggested large differences in performance within the group of HD participants for trustworthiness and dominance perception, in that some HD participants seemed to be more impaired on the trustworthiness and dominance tasks than other HD group members.

To objectively verify this impression, we ran a two-step cluster analysis on the Fisher r to z transformed values of the trustworthiness and dominance perception measures and their associated gender perception comparison measures. For this, we included all 67 participants (41 HD participants, and 26 controls) who took part in the study.

This resulted in a two-cluster solution, with one cluster consisting of especially impaired HD participants (n=10), and one cluster consisting of control and less impaired HD participants (31 HD participants, and 26 controls). Figures 3c and 3d show performance of the trustworthiness and dominance perception tasks broken down into these HD subgroups.

It would of course be circular to analyse the data shown in Figures 3c and 3d; they are presented here simply to demonstrate that the cluster analysis did create a clear separation between subgroups of HD participants with relatively impaired or unimpaired trustworthiness and dominance perception. It should be emphasised, though that these subgroups can then be used to create meaningful (i.e. not circular) comparisons across other measures; the subgroups are defined on the measures of trustworthiness and dominance perception, but we will compare the resulting subgroups on other measures of performance.

The partitioning of HD participants into subgroups based on cluster analysis of the trustworthiness and dominance perception tasks allows, as a complement to the comparison of controls and the entire HD group, a more refined analysis strategy for the other measures included in our study. Therefore, for the Benton Facial Recognition Test, the Mind in the Eyes Test, and the Ekman 60 Faces Test, we will 1.) report results of the comparison of controls and the full HD participant group and 2.), results of the comparison of the neurologically normal control group, and of the subgroups of HD participants with relatively impaired or unimpaired trustworthiness and dominance perception. Later on, we will also use the subgroups to explore brain microstructural integrity as measured with DTI.

In order to investigate both behavioural and neurological (DTI) measures, we compared performance of the control group and the unimpaired and impaired HD subgroups, excluding nine HD participants who did not have DTI measures (6 in the unimpaired and 3 in the impaired HD group). This resulted in 3 unequally sized groups - a group of 25 unimpaired HD participants, a group of 7 impaired HD participants, and a group of 26 controls.

Table 1b about here, please

Since the 'impaired HD' subgroup is very small, we decided to use Bonferroni corrected non-parametric procedures in statistical analyses involving this group to counter possible violations of normality. The level of significance involving 3 comparisons on each measure will therefore be $p = .016$.

Clinical and demographic details for the two HD subgroups for which DTI measures were taken are given in Table 1b. The HD subgroups did not differ from controls with respect to age (Kruskal-Wallis $\chi^2(2) = 1.46$, $p = .48$), years of formal education (Kruskal-Wallis $\chi^2(2) = 2.85$, $p = .24$), and IQ (Kruskal-Wallis $\chi^2(2) = 4.69$, $p = .096$).

There was some evidence that the HD subgroup with impaired performance of the social perception tasks were at a more advanced stage of disease progression. The UHDRS Motor score (Mann-Whitney $U = 26.0$, $p = .008$, $r = .38$) was significantly different between the HD subgroups, and for the UHDRS TFC (Mann-Whitney $U = 36.0$, $p < .037$, $r = .38$), disease duration (Mann-Whitney $U = 39.0$, $p = .071$, $r = .33$) and disease burden (Mann-Whitney $U = 36.0$, $p = .037$, $r = .35$) there was a trend towards statistical significance. Between the two HD subgroups, there were no significant differences in respect to the number of CAG repeats on the long allele (Mann-Whitney $U = 56.5$, $p = .38$, $r = 0.16$), and the number of CAG repeats on the short allele (Mann-Whitney $U = 55.5$, $p = .35$, $r = .17$).

Face perception and social cognition tasks

Benton Facial Recognition Test and Mind in the Eyes Test: For the overall group comparison, Student's t-tests showed significant differences between participants with HD and controls, reflecting poorer overall performance by HD participants on the Benton Facial Recognition Test ($t(65) = 3.50$, $p = .001$, $d = 0.89$) and the Mind in the Eyes Test ($t(65) = 3.11$, $p = .003$, $d = 0.76$). See Table 2 for details.

For the comparisons involving controls and the two HD subgroups, Kruskal-Wallis tests showed significant differences between groups on the Benton Facial Recognition Test (Kruskal-Wallis $\chi^2(2) = 9.58$, $p = .008$) and the Mind in the Eyes Test (Kruskal-Wallis $\chi^2(2) = 11.6$, $p = .003$).

Subsequent Mann-Whitney tests showed a significant deficit on the Benton test for the unimpaired HD subgroup when compared with the control group (Mann-Whitney $U = 175.0$, $p = .005$, $r = .51$). A trend towards significance was found for the difference between the control group and the impaired HD subgroup (Mann-Whitney $U = 45.5$, $p = .043$, $r = .36$), and no significant difference between both HD subgroups (Mann-Whitney $U = 74.5$, $p = .56$, $r = .11$). Although lower than that of controls, the performance scores of both HD subgroups was in a range considered normal by the test manual.

For the Mind in the Eyes Test, Mann-Whitney tests showed a significant deficit for the impaired HD subgroup, when compared with controls (Mann-Whitney $U = 26.0$, $p = .003$, $r = .51$), a trend towards significance for the unimpaired HD subgroup when compared to controls controls (Mann-Whitney $U = 218.0$, $p = .043$, $r = .35$), and a significant difference between the impaired and unimpaired HD subgroup (Mann-Whitney $U = 33.0$, $p = .013$, $r = .44$). On this test, which assesses the ability to recognise emotions and mood states from the faces of others, the impaired HD sub-group identified from the trustworthiness and dominance perception tasks was disproportionately impaired compared to the other groups.

Table 2 about here, please

Ekman 60 Faces Test: To evaluate overall performance of the HD and control groups we performed a mixed ANOVA with emotion (happiness, surprise, fear, sadness, disgust, and anger) as a within-subject factor, and group (HD and controls) as a between-group factor. This showed a significant effect of emotion ($F(5,325) = 54.04$, $p < .001$, $\eta p^2 = .45$), a significant effect of group ($F(1,65) = 25.1$, $p < .001$, $\eta p^2 = .28$) based on poorer performance by HD than control participants, and a significant emotion x group interaction ($F(5,325) = 4.80$, $p < .001$, $\eta p^2 = .069$).

To explore the emotion x group interaction in more detail, we performed t-tests across the six basic emotions; this gave a non-significant result for happiness ($t(65) = 0.99$, $p = .32$, $d = 0.28$), but significant results reflecting deficits in the HD group for recognising surprise ($t(65) = 2.25$, $p = .028$, $d = 0.58$), fear ($t(65) = 2.48$, $p = .01$, $d = 0.62$), and sadness ($t(65) = 3.04$, $p = .003$, $d = 0.76$). Overall, though, disgust ($t(62.7) = 4.50$, $p < .001$, $d = 1.20$) and anger ($t(65) = 5.46$, $p < .001$, $d = 1.42$) were the most severely impaired emotions for HD participants. See Figure 4a for details.

Figure 4a and 4b about here, please

For the subgroups comparison, we used a Kruskal-Wallis test to analyse the average number of correct responses across all emotions. These were for the controls: M = 46.3, SD = 7.31, - for the unimpaired HD subgroup: M = 40.1, SD = 8.20, - and for the impaired HD subgroup: M = 25.7, SD = 7.80. There was a significant difference between groups (Kruskal-Wallis $\chi^2(2) = 19.5$, $p < .001$).

Subsequently performed Mann-Whitney test decomposed the significant group difference and showed a significant difference between the better performing controls and both the unimpaired (Mann-Whitney $U = 187.0$, $p = .009$, $r = .46$) and impaired HD group (Mann-Whitney $U = 7.50$, $p < .001$, $r = .65$). In addition, the unimpaired HD group performed significantly better than the impaired HD group on this test (Mann-Whitney $U = 14.5$, $p < .001$, $r = .59$).

To see which emotions are particularly affected, we performed Kruskal-Wallis tests; this gave significant results for happiness (Kruskal-Wallis $\chi^2(2) = 8.46$, $p = .015$), surprise (Kruskal-Wallis $\chi^2(2) = 15.5$, $p = .001$), sadness (Kruskal-Wallis $\chi^2(2) = 11.9$, $p = .003$), disgust (Kruskal-Wallis $\chi^2(2) = 13.2$, $p = .001$)., and anger (Kruskal-Wallis $\chi^2(2) = 20.5$, $p < .001$). There was no significant group difference for fear (Kruskal-Wallis $\chi^2(2) = 5.96$, $p = .051$). To explore the group differences across emotions in more detail, we used Mann-Whitney tests.

Subsequent tests showed no significant differences between the control group and the unimpaired HD subgroup for recognising happiness, surprise, fear, and sadness ($p > .12$). However, the recognition rates for disgust showed a trend towards significance (Mann-Whitney $U = 202.5$, $p = .019$, $r = .42$) and there was a significant impairment for anger (Mann-Whitney $U = 155.0$, $p = .001$, $r = .47$) . There was a significant difference between the control group and the impaired HD subgroup for surprise (Mann-Whitney $U = 12.5$, $p < .001$, $r = .40$), sadness(Mann-Whitney $U = 21.5$, $p = .001$, $r = .55$), disgust (Mann-Whitney $U = 14.5$, $p < .001$, $r = .60$) , and anger (Mann-Whitney $U = 10.0$, $p < .001$, $r = .65$) , but not for happiness (Mann-Whitney $U = 51.0$, $p = .082$, $r = .42$) , there was a trend towards significance for fear (Mann-Whitney $U = 40.0$, $p = .024$, $r = .40$) . The recognition rate for surprise (Mann-Whitney $U = 13.5$, $p < .001$, $r = .61$), sadness (Mann-Whitney $U = 21.5$, $p = .001$, $r = .54$), and anger (Mann-Whitney $U = 28.0$, $p = .005$, $r = .48$) was significantly lower for the impaired compared to the unimpaired HD subgroup. No significant differences for the HD subgroups were found for fear (Mann-Whitney $U = 60.0$, p

= .23, *r* = .22) , disgust (Mann-Whitney *U* = 50.5, *p* = .092, *r* = .30), and happiness(Mann-Whitney *U* = 45.0, *p* = .054, *r* = .48). On this test, then, the unimpaired HD subgroup showed mild deficits in emotion recognition, while the impaired HD subgroup showed substantial deficits in emotion recognition across a wide range of facial expressions. See Figure 4b for details.

Correlations between measures of face perception and social cognition

To explore the relation between impairments affecting face perception and social cognition, and especially to see whether it is plausible that accurate facial expression recognition might be a prerequisite for perceiving trustworthiness and dominance in the faces of others, we correlated performance of all HD participants on the trustworthiness and dominance perception tasks with their performance on the Benton Facial Recognition Test, the Ekman 60 Faces Test, and the Mind in the Eyes Test. No correlation was found with the Benton Test, but there were significant correlations with the Ekman 60 and the Mind in the Eyes Tests, both of which involve interpreting expressions of moods and feelings, either across the whole face (Ekman 60 Faces Test) or from the eye region alone (Mind in the Eyes Test). These correlations were usually stronger for the trustworthiness and dominance rating tasks than for the femininity rating tasks. See Table 3 for details. On this basis, we can conclude that there is a plausible initial case for a common underlying factor involving the interpretation of facial expressions and first impressions of social traits.

Table 3 about here, please

Diffusion Tensor Imaging (DTI)

As part of our analysis strategy, we identified HD subgroups based on their relatively intact or impaired performance of the trustworthiness and dominance perception tasks and then compared the performance of these subgroups on other behavioural tests of face perception. To complement the differences found using these behavioural measures, we used DTI to examine white matter organisation as a sensitive index of subtle regional neural degeneration. To this end, the group of participants with HD for whom DTI data were available was divided into two subgroups based on the results of the cluster analysis of the trustworthiness and dominance perception tasks. Whole-brain based spatial t-statistics of the DTI data revealed 15 clusters with a voxel size of greater than 268 (see DTI methods), in which FA was significantly reduced in the subgroup showing impaired performance on the trustworthiness and dominance tasks compared to the unimpaired HD subgroup. Two clusters with less than 268 voxels located in the insula cortex which can be regarded as extensions of larger clusters in this region (cluster 16 and 17, see Table 4) are also reported here.

FA was decreased in the corpus callosum, the ACC and other fronto-medial regions, the amygdalar-insular complex bilaterally, the cerebellum and the pons (see Table 4 for MNI coordinates, cluster size and mean significance levels of the identified clusters.

To see whether the identified structures were particularly associated with aspects of social cognition and not with cognition in general, we ran a univariate ANOVA with averaged FA values of identified clusters as the dependent variable and group ('impaired HD group' vs. 'unimpaired HD group') as a fixed factor. This gave a significant group effect ($F(1,32) = 23.1$, $p < .001$, $\eta p^2 = .71$). We included a number of cognitive measures (routinely collected for all people with HD as part of the UHDRS, and not reported here in detail) such as lexical ($F(1,32) = 0.002$, $p = .97$, $\eta p^2 < .001$) and semantic word fluency ($F(1,32) = 0.55$, $p = .47$, $\eta p^2 = .021$), performance on the Symbol Digit Modality Test ($F(1,32) = 1.62$, $p = .21$, $\eta p^2 = .059$), and the Stroop Test ($F(1,32) = 0.29$, $p = .59$, $\eta p^2 = .011$) as covariates. None of these covariates was significant. This suggests that the identified regions are more likely associated with aspects of social cognition then other forms of information processing.

Table 4 about here, please

We next assessed whether there was an association of the overall mean FA values of brain regions associated with recognition of trustworthiness and dominance with indicators of HD severity. There was a significant correlation of FA values with the UHDRS motor score ($r(32) = -.57$, $p = .001$) and disease burden ($r(32) = .49$, $p = .008$).

To explore the potential role of facial emotion recognition in recognising trustworthiness and dominance, we correlated the FA values of the identified cerebral regions associated with impaired trustworthiness and dominance perception with the performance measures from the Mind in the Eyes Test and the Ekman 60 Faces Test. Here we found overall strong correlations (see Table 5 for details).

Table 5 about here, please

To summarise, DTI identified neural structures predominantly associated with trust and dominance evaluation but not with performance on standard neuropsychological background tests. FA values of these structures correlated with the ability to recognise emotions as assessed with two independent tasks. These findings complement those from the behavioural measures and further support the initial case for a common underlying factor for interpretation of facial expressions and first impressions of social traits.

DISCUSSION

Our overarching aim was to investigate the ability to form facial first impressions of trustworthiness and dominance in people with HD. In this respect, the striking finding is that HD participants could be objectively divided into subgroups with relatively intact or relatively impaired perception of trustworthiness and dominance. We were then able to look at other behavioural and neural correlates of this overall difference.

First, though, we need to point out the main features of the novel method we developed for testing perception of trustworthiness and dominance. This used computer image manipulation to create quasi-linear continua of face-like images which varied in perceived trustworthiness or dominance, allowing us to measure how well these subtle gradations in manipulated trait impressions were perceived. These continua of face-like images were created from prototypes that represented the high or low trustworthiness or dominance of male and female faces through averaging a number of suitably-rated everyday photographs drawn from a large set of 1,000 images chosen to represent the variability of faces seen in everyday life. This empirical approach of averaging everyday photographs allows us to capture all the cues that are consistent across a set of images without predetermining what these cues might be. This is important in circumstances where perceivers may base their first impression on multiple cues whose roles are not well understood. For example, Figure 2 shows that the cues that signal changes in perceived trustworthiness include a combination of changes in age, face shape, skin tone, and facial expression.

To help us interpret findings from these trustworthiness and dominance perception tasks we created comparable measures of the perception of the structural characteristic of face gender (i.e., the degree of masculine or feminine appearance). This allowed us to estimate the extent to which problems in perception of any relatively subtle facial characteristic might spill over to the evaluation of trustworthiness or dominance from facial appearance.

We found significantly impaired performance in evaluating trustworthiness and dominance for the overall HD group when compared with controls, but both groups performed more similarly on the femininity rating tasks. However, cluster analysis demonstrated that only a subgroup of HD participants showed a severe impairment of both trustworthiness and dominance perception. This problem therefore encompassed both dimensions of Oosterhof and Todorov's (2008) model, showing that it is a pervasive deficit in social trait evaluation. Importantly, though, participants from the HD subgroup with impaired perception of trustworthiness and dominance showed only minor (or no) problems in determining the gender-stereotypicality of faces and were still in the normal range of the Benton Facial Recognition Test (even though they performed this task less well than normal participants), showing that they remain able to evaluate quite subtle changes in appearance. In other words, this is not a perceptual deficit as such, it is a problem in evaluating social significance.

Because we also tested ability to recognise facial expressions of emotion, these problems in perceiving trustworthiness and dominance could be used to explore the plausibility of the emotion overgeneralisation hypothesis. On the Mind in Eyes Test, which includes some items testing emotion recognition, only the impaired HD subgroup differed significantly from controls, while the unimpaired HD subgroup was largely unaffected. A sensitive measure of emotion recognition deficits is the Ekman 60 Faces Test. This test showed that the overall HD group were impaired in facial emotion recognition, affecting all basic emotions except happiness. Recognition of disgust and anger was particularly affected, mirroring results from previous studies (Calder, et al., 2010; Sprengelmeyer, et al., 1996). This may be of interest because trustworthiness ratings of faces by neurologically normal perceivers are particularly low when the faces show expressions of disgust and anger (Willis et al., 2011; Sutherland et al., in press). However, looking at the HD subgroups in more detail, we only found differentially severe deficits in disgust and anger recognition in the unimpaired group, with instead a severe deficit in recognising all basic emotions except happiness and fear in the impaired subgroup.

We also found clear correlations between impairments affecting the evaluation of trustworthiness and dominance and problems in recognising emotion based on the Ekman 60 and Mind in the Eyes Tests. These correlations would be expected from Montepare and Dobish's (2003) emotion overgeneralisation hypothesis. They are at first sight less obviously consistent with Oosterhof & Todorov's (2008) model, in which the dominance dimension is suggested to particularly involve cues to physical strength, masculinity and maturity (Keating, 1985; Oosterhof & Todorov, 2008). However, this idea does not rule out a possible influence of facial expression alongside these more structural facial characteristics, and it is important to note that by creating a matrix of images for this study in which perceived dominance and gender were as far as possible orthogonally manipulated, we may well have reduced the impact of gender typicality on the evaluation of the dominance dimension. This may explain why the correlations of the trustworthiness stimuli with the Mind in the Eyes and Ekman 60 faces tasks showed a clearer separation between the trustworthiness and femininity ratings than was evident for the stimuli used to evaluate perception of dominance (see Table 3). Moreover, we need to bear in mind that for reasons explained in the Method section our new social perception tasks were given in a fixed order, and in consequence there may have been some fatigue or 'carry-over' effects (Rhodes, 1986) on the ratings of the stimuli from the dominance x gender matrix.

Our findings from purely behavioural data show that the impairments found in HD can enhance our understanding of mechanisms involved in normal social cognition.

However, we were also able to complement these behavioural data with measures based on DTI. By contrasting regional FA between subgroups of HD participants showing impairments in judging trustworthiness and dominance with the FA measures of HD participants who were able to evaluate these social dimensions, we could provisionally identify candidate neural structures associated with social cognition. Regions whose FA was significantly decreased in the impaired HD group compared to the unimpaired group included the right dorsolateral and left medial frontal cortex, the right ACC, the corpus callosum, the left somatosensory cortex and the right fusiform gyrus. The anterior and posterior portions of both the left and right insular cortex and the left and right amygdalar-parahippocampal region were also prominently affected. In addition, we found significant decreases in FA for the impaired HD subgroup in the left and right cerebellum, and the pons.

Other studies associate evaluation of facial trustworthiness with the amygdala (Adolphs, Tranel, & Damasio, 1998; Engell, Haxby, & Todorov, 2007; Hall, et al., 2010; Mattavelli, et al., 2012; Todorov, 2008; Todorov & Engell, 2008; Winston, et al., 2002), the fusiform gyrus, and the insular cortex (Winston, et al., 2002). Moreover, the fusiform gyrus, the superior temporal gyrus and the lingual gyrus are associated with evaluation of dominance from facial appearance and dominant head postures (Chiao, et al., 2008).

It is important to know whether these identified regions are primarily associated with aspects of social cognition or with cognition in general. We therefore compared averaged FA values of the impaired and unimpaired HD subgroups with a set of cognitive performance measures as covariates, all of which were found to be non-significant. This suggests that the identified clusters are predominantly involved in social cognition. Taken together, then, regions involved in evaluating trustworthiness and dominance identified in our HD study mirror previous findings and map quite consistently the boundaries of the social brain (Adolphs, 2009).

We do not know for certain why regions of the social brain should be differentially severely affected in a subgroup of people with HD, but inspection of the clinical characteristics show that the impaired HD group was in a slightly more advanced stage of the disease than the unimpaired group, with 2.3 compared to 4.4 years of disease duration, respectively. A recent study by McColgan and co-workers (McColgan et al., 2015) found that specific functionally defined networks in the brains of people with HD, defined by highly interconnected regions working as a strongly interlinked collective (van den Heuvel and Sporns, 2011), may become dysfunctional because they lose their connectivity during progression of the disease. This could be shown for networks involved in motor and cognitive functions in HD (Poudel et al., 2014, McColgan et al., 2015). Some of these networks lose connectivity in the pre-symptomatic stages of the disorder, while other networks are affected after occurrence of the first motor signs (Poudel et al., 2014).

If the driving force of the neuropathology of HD is the breakdown of neural networks, one could argue that with progression of the disease one network after another, depending on the network's vulnerability and resilience, may become dysfunctional, possibly in a predictable order. It may therefore well be that the network underlying social cognition – the social brain - is affected in HD at a certain point of disease development, potentially representing a milestone of disease progression with a characteristic combination of symptoms. Because the break down of social cognition has implications for the understanding and clinical management of HD, further longitudinal studies looking at a variety of social cognitive dimensions are warranted.

To further address the possible link between emotion recognition and trustworthiness and dominance evaluation as suggested by the emotion overgeneralisation hypothesis (Montepare & Dobish, 2003), we correlated mean FA values of the identified regions with the HD participants' performance scores on the Mind in the Eyes Test and the Ekman 60 Faces Test. Here we found high correlations between changes in regional FA values and behavioural measures of emotion recognition, which supports the already established behavioural link between emotion recognition and trustworthiness and dominance evaluation through data at the neuroanatomical level.

A note of caution is necessary, though, since although correlations are a prerequisite of a causal link, they do not in themselves establish a direct causal connection. It may well be the case that the deficit in emotion recognition is not itself a prerequisite for impaired trustworthiness and dominance evaluation, but rather a more general consequence of a disintegrating social brain. If this is correct, and the social brain of the impaired HD subgroup is more or less uniformly affected by a pathological process, we do not know whether some of the identified neural structures might be functionally more closely related to evaluation of trustworthiness and dominance, whilst others might be associated with emotion recognition per se. Indeed, a different pattern of results has been reported by Willis and co-workers (Willis, Palermo, Burke, McGrillen, & Miller, 2010). They found that patients with orbitofrontal cortex lesions showed differences from controls in judging the approachability of emotional faces but did not have an overall deficit in facial expression recognition. As well as the difference in participants (HD vs OFC lesions), the procedures of Willis and co-workers differ from ours in terms of the dimensions investigated (though approachability is closely related to trustworthiness), the stimuli evaluated (computer-manipulated dimensions vs category-based photographs) and the facial expression recognition tests used (Ekman vs KDEF faces). None the less, their findings suggest that some neural structures might be functionally more closely related to evaluation of social significance, whilst others might be associated with emotion recognition per se. One possible reason may be that OFC is known to be involved in linking stimuli to rewards and emotion-related learning (O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001) and this role may be emphasised through procedures (such as Willis et al., 2010) which involve evaluating the social significance of highly emotional faces. Alternatively, a later study by Willis, Palermo, McGrillen and Miller (2014) found that OFC lesions did create problems in more demanding tests of facial expression recognition. This makes it seem likely that trait judgements to emotional faces could turn out to be a more sensitive measure of social cognitive capacities than emotion labelling.

We have shown, then, that the impairments found in HD can enhance our understanding of mechanisms involved in normal social cognition. Further detailed studies of HD and more focal impairments of neural components of the 'social brain' offer considerable promise for clarifying how we as humans form our first impressions of others.

Acknowledgements

The authors thank all participants who took part in this study.

FIGURE LEGENDS

Figure 1a and 1b: The left panel (Figure 1a) shows the Trustworthiness x Gender Matrix. Each of the 10 rows shows from left to right the incremental increase from non-trustworthiness to trustworthiness. The 10 columns show from top to bottom the incremental change from strongly male to strongly female appearance. The right panel (Figure 1b) shows the Dominance x Gender matrix. Each of the 10 rows shows from left to right the incremental increase from non-dominance to dominance. The 10 columns show from top to bottom the incremental change from strongly male to strongly female appearance. Within each matrix, the four images marked with a red border are computer-averaged prototypes representing relatively high and low levels on each dimension. The rest of the 10x10 matrix is then created by interpolating morphed and caricatured images around these prototypes in the horizontal and vertical dimensions.

Figure 2a to 2d: Examples of stimulus material used in the Trustworthiness and Dominance tasks, Trustworthiness rating (Figure 2a), Femininity rating based on an image from the Trustworthiness x Gender matrix (Figure 2b), Dominance rating (Figure 2c), Femininity rating based on an image from the Dominance x Gender matrix (Figure 2d).

Figure 3a to 3d: Performance of control participants and all HD participants on the Trustworthiness (Figure 3a) and Dominance (Figure 3b) rating tasks, with performance expressed as the mean correlation between the level of computer generated social trait intensity or femininity (1 to 10) and the participant's mean rating of each level of trait intensity or femininity, respectively. Figures 3c and 3d show performance of control participants and HD participants with and without impaired trustworthiness and dominance recognition; these are included to demonstrate that the cluster analysis achieved an effective separation between normally-performing and impaired HD subgroups.

Figure 4a and 4b: Results from the Ekman 60 Faces test. Figure 4a shows performance of control participants and all participants with HD. Figure 4b shows performance of control participants and HD participants with and without impaired trustworthiness and dominance recognition.

Figure 1

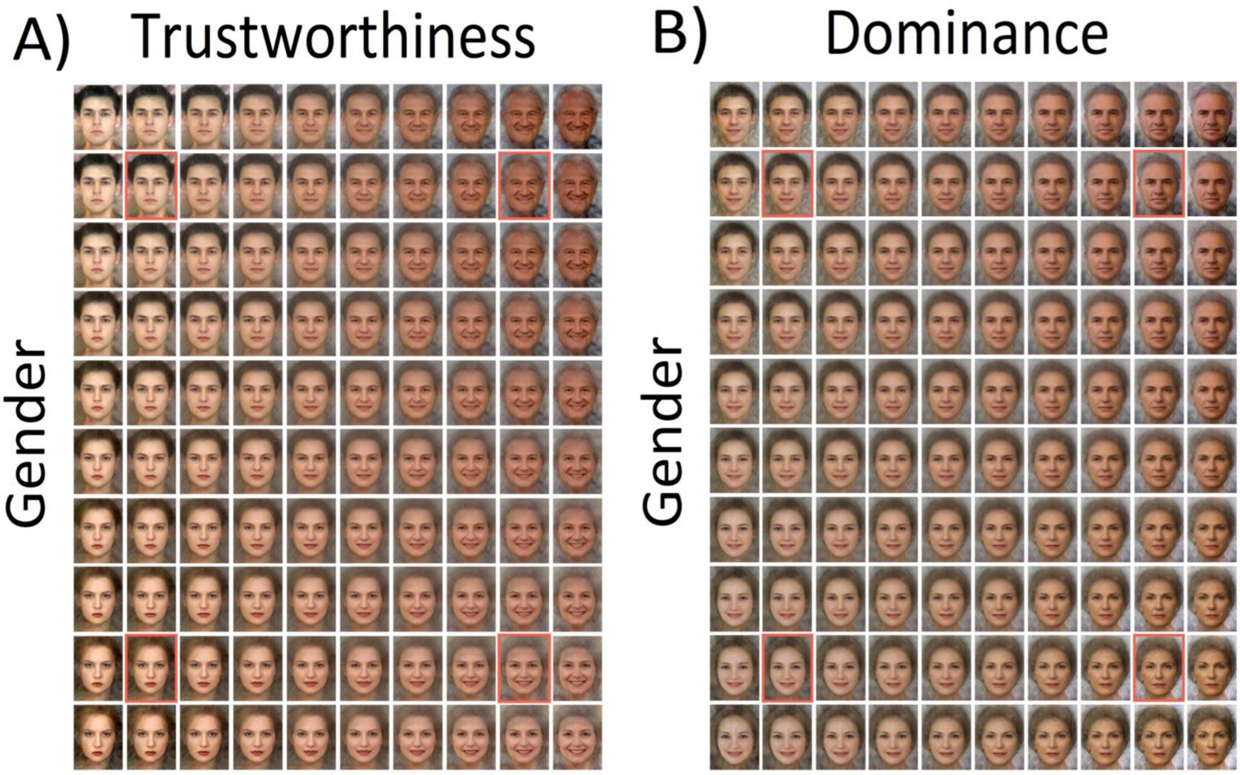

Figure 2

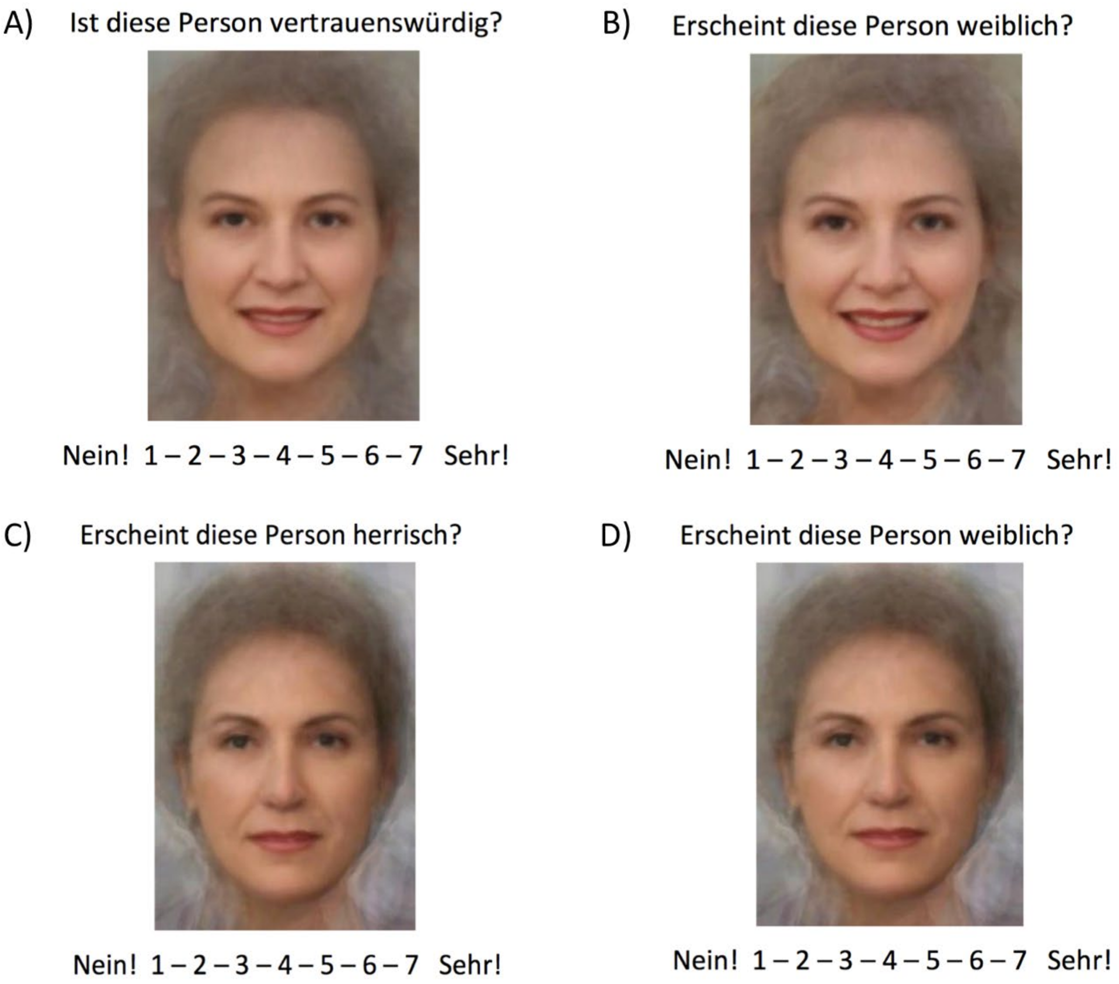
A)
Ist diese Person vertrauenswürdig?
Nein! 1 – 2 – 3 – 4 – 5 – 6 – 7 Sehr!
B)
Erscheint diese Person weiblich?
Nein! 1 – 2 – 3 – 4 – 5 – 6 – 7 Sehr!
C)
Erscheint diese Person herrisch?
Nein! 1 – 2 – 3 – 4 – 5 – 6 – 7 Sehr!
D)
Erscheint diese Person weiblich?
Nein! 1 – 2 – 3 – 4 – 5 – 6 – 7 Sehr!

Figure 3

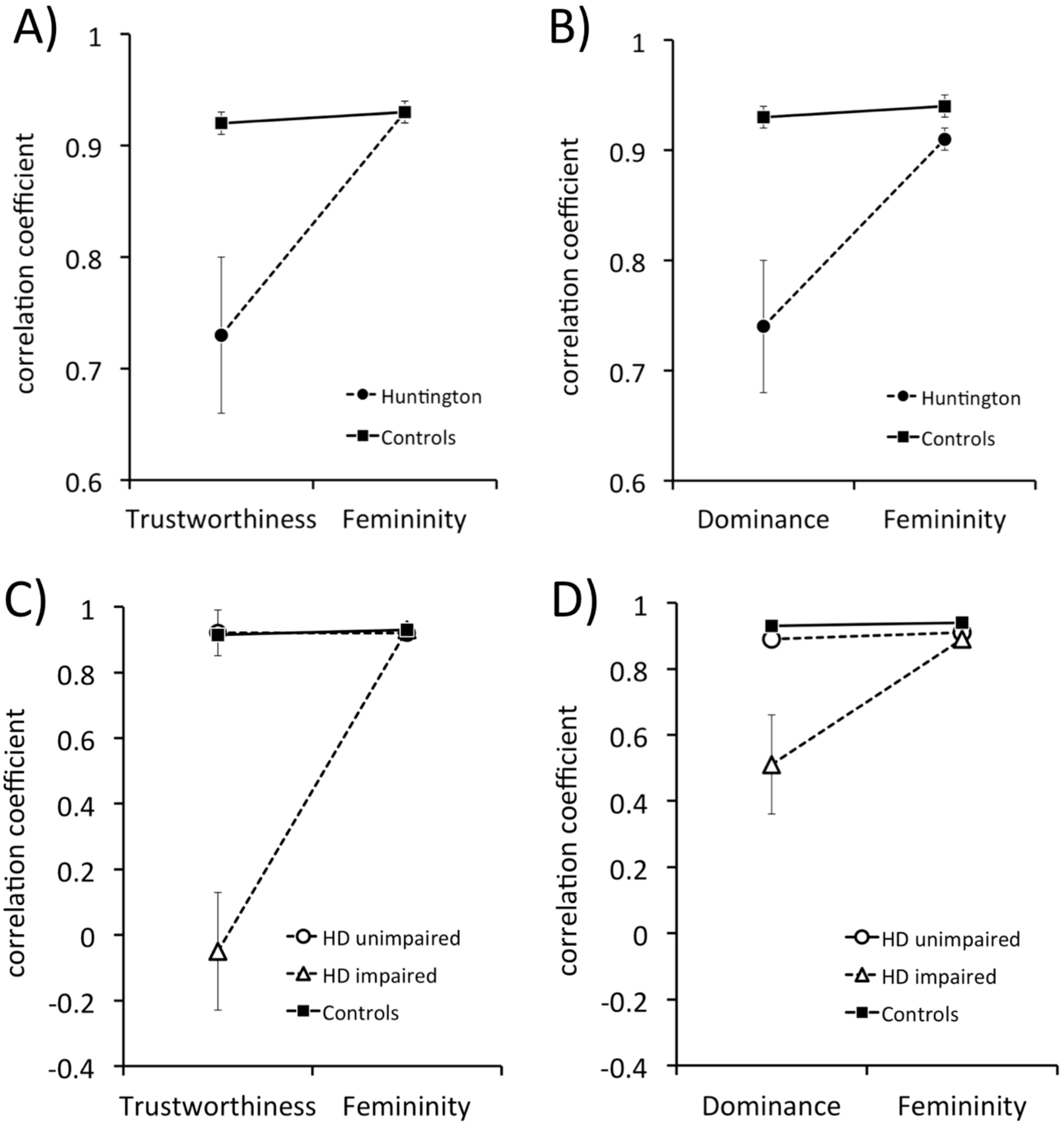
A)
correlation coefficient
1
0.9
0.8
0.7
0.6
Trustworthiness
Femininity
Huntington
Controls
B)
correlation coefficient
1
0.9
0.8
0.7
0.6
Dominance
Femininity
Huntington
Controls
C)
correlation coefficient
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
Trustworthiness
Femininity
HD unimpaired
HD impaired
Controls
D)
correlation coefficient
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
Dominance
Femininity
HD unimpaired
HD impaired
Controls

Figure 4

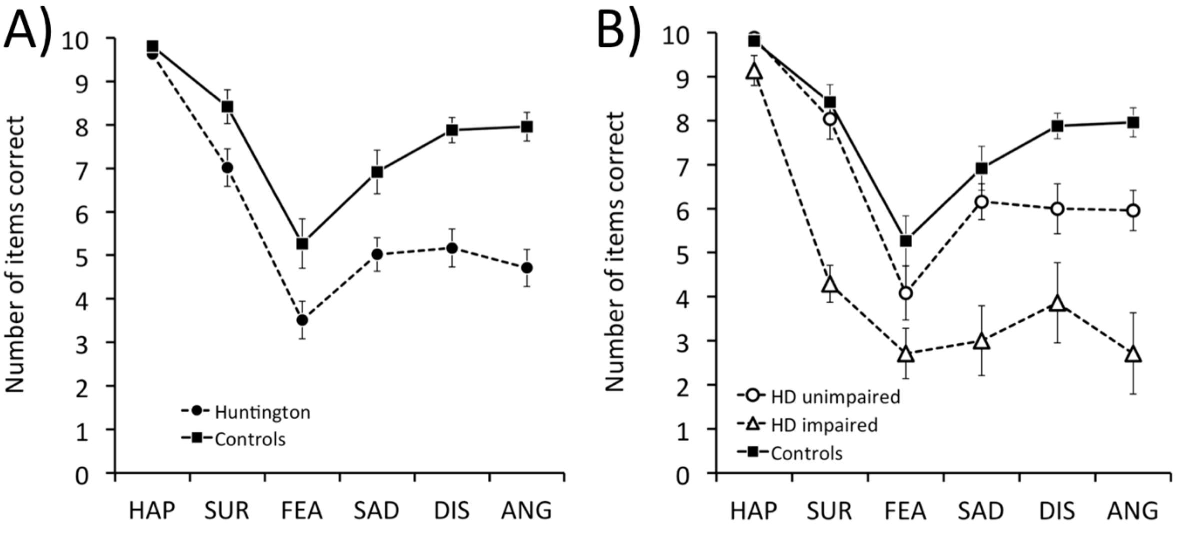
A)
Number of items correct
10
9
8
7
6
5
4
3
2
1
0
Huntington
Controls
HAP
SUR
FEA
SAD
DIS
ANG
B)
Number of items correct
HD unimpaired
HD impaired
Controls
HAP
SUR
FEA
SAD
DIS
ANG

Table 1a Background information for neurologically normal and HD participants.

	Controls		HD	
	Mean	(SD)	mean	(SD)
Number of participants	n = 26		n = 41	
Age (years)	47.0	(9.5)	48.7	(10.0)
Education (years)	14.7	(2.7)	13.9	(2.5)
IQ	104.5	(13.3)	100.1	(10.9)
UHDRS motor			21.3	(13.7)
UHDRS TFC			10.9	(2.3)
CAG large			43.0	(2.2)
CAG short			18.6	(2.9)
Disease Burden*			356.8	(88.6)
Time since disease onset (years)			4.3	(4.8)

* Disease Burden is calculated: [(CAG Allele large – 35.5)*age] (Penney, Vonsattel, MacDonald, Gusella, & Myers, 1997).

Table 1b Background information for neurologically normal participants, and for the HD subgroups without and with impaired perception of trustworthiness and dominance for which DTI data were available.

	Controls		HD non-imp		HD imp.	
	Mean	(SD)	mean	(SD)	mean	(SD)
Number of participants	n = 26		n = 25		n = 7	
Age (years)	47.0	(9.5)	45.8	(8.6)	50.7	(10.1)
Education (years)	14.7	(2.7)	14.4	(2.7)	12.9	(1.5)
IQ	104.5	(13.3)	103.4	(11.4)	94.3	(6.0)
UHDRS motor			14.6	(10.4)	29.9	(13.6)
UHDRS TFC			12.1	(1.1)	10.7	(1.5)
CAG large			42.7	(2.3)	43.6	(2.1)
CAG short			18.3	(3.1)	20.3	(3.2)
Disease Burden*			323.3	(88.5)	394.2	(55.6)
Time since disease onset (years)			2.3	(2.6)	4.4	(3.5)

* Disease Burden is calculated: [(CAG Allele large – 35.5)*age] (Penney, et al., 1997).

Table 2 Performance of the Benton Facial Recognition and the Mind in the Eyes Test for neurologically normal and HD participants, and for the HD subgroups without and with impaired perception of trustworthiness and dominance for which DTI data were available.

	Controls		HD		HD non-imp.		HD imp.	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Number of participants	n = 26		n = 41		n = 25		n = 7	
Benton Facial Recognition	46.9	(6.0)	41.4	(6.3)	41.6	(7.1)	41.1	(5.6)
Mind in the Eyes Test	23.3	(6.1)	19.1	(5.0)	20.7	(3.8)	15.6	(4.2)

Table 3 Pearson correlations between performance of the Benton Facial Recognition test, the Mind in the Eyes Test, and the Ekman 60 Faces Test with Fisher-transformed performance on the Trustworthiness and Dominance rating tasks and the gender rating comparison tasks for HD participants (n=41).

	Benton	Mind in the Eyes	Ekman 60 Faces
Trustworthiness rating	.03	.48 **	.54 ***
Femininity rating of trustworthiness images	-.09	.31	.30
Dominance rating	.15	.44 *	.48 **
Femininity rating of dominance images	.17	.44 *	.37 *

* = p<.05; ** = p<.01; *** = p<.005

Table 4 Cerebral regions with reduced fractional anisotropy (FA) in HD participants with impaired trustworthiness and dominance recognition.

No.[1]	voxel	x	y	z	Anatomical region
1	3207	0	-10	25	Corpus Callosum
2	2960	-20	8	48	Left Medial Frontal Gyrus
7	620	16	38	-3	Right ACC
3	2394	30	4	44	Right middle frontal Gyrus
Occipital and Parietal Lobe					
4	1389	33	-73	10	Right Fusiform Gyrus
9	588	-27	-24	44	Left Postcentral Gyrus
Insula and Amygdala Region					
5	1116	-46	-28	1	Left Posterior Insula
17	252	-37	1	-9	Left Anterior Insula
6	982	40	-17	-7	Right Posterior Insula
15	269	28	-40	23	Right Posterior Insula
16	262	36	1	-9	Right Anterior Insula
8	599	32	28	3	Right Anterior Insula
13	397	-14	-9	-10	Left Amygdala – Parahippocampal Region
10	428	14	-15	-10	Right Amygdala - Parahippocampal Region
Cerebellum/Brainstem					
12	401	-11	-55	-21	Cerebellum/Left anterior Lobe
11	427	5	-43	-30	Cerebellum/Right Anterior Lobe
14	333	-2	-24	-19	Pons

---

[1] No. = cluster rank based on voxel size from 1 = largest to 17 smallest

Table 5 Correlation of the overall HD group's performance on the Mind in the Eyes Test (MiE) and the Ekman 60 Faces Test (EK60) (scores for recognising each emotion, and total score) with the HD group's FA values (n = 31) of cerebral regions identified in HD participants with impaired trustworthiness and dominance recognition.

		MiE	EK60			EK60			
No.[1]	Anatomical region		Total	Hap	Sur	Fea	Sad	Dis	Ang
1	Corpus Callosum	**.74*** **	**.64*** **	.28	.47**	.41*	.46**	.36*	.67***
2	Left Medial Frontal Gyrus	**.56** **	**.59*** **	.20	.54**	.39*	.30	.46**	.44*
7	Right ACC	**.72*** **	**.64*** **	.27	.44*	.43*	.47**	.46**	.55**
3	Right middle frontal Gyrus	**.63*** **	**.64*** **	.25	.50**	.37*	.45**	.53**	.50**
Occipital and Parietal Lobe									
9	Left Postcentral Gyrus	**.65*** **	**.57** **	.08	.39*	.41*	.43*	.45**	.43*
4	Right Fusiform Gyrus	**.68*** **	**.62*** **	.12	.48**	.36*	.38	.51**	.55**
Insula and Amygdala Region									
5	Left Posterior Insula	**.61*** **	**.53** **	.18	.51**	.32	.33	.38*	.39*
17	Left Anterior Insula	**.44* **	**.54** **	.51**	.36*	.32	.40*	.32	.53**
6	Right Posterior Insula	**.56** **	**.72*** **	.32	.53**	.38*	.52**	.59***	.62***
15	Right Posterior Insula	**.44* **	**.58*** **	.32	.58*	.07	.37*	.64***	.46**
16	Right Anterior Insula	**.63*** **	**.67*** **	.32	.47**	.46**	.53**	.38*	.62***
8	Right Anterior Insula	**.61*** **	**.46** **	.15	.30	.35*	.30	.33	.40*
13	Left Amygdala/ Parahippocampal Region	**.29**	**.54** **	.24	.27	.41*	.42*	.42*	.46**
10	Right Amygdala/ Parahippocampal Region	**.32**	**.49** **	.15	.32	.30	.29	.47**	.42*
Cerebellum/Brainstem									
12	Left Cerebellum	**.44* **	**.50** **	.17	.40*	.23	.39*	.49**	.32
11	Right Cerebellum	**.45* **	**.43* **	.30	.23	.48**	.30	.16	.37
14	Pons	**-.41* **	**-.41* **	-.34	-29	-.27	-.42*	-.27	-.24

[1] No. = cluster rank based on voxel size from 1 = largest to 17 smallest

* = p<.05; ** = p<.01; *** = p<.005

References

Adams, R. B., Jr., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: a mechanism for impression formation? *Cogn Emot, 26*, 431-441.

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology, 60*, 693-716.

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature, 393*, 470-474.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 241-251.

Basser, P. J., & Jones, D. K. (2002). Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR in Biomedicine, 15*, 456-467.

Benton, A. L., Sivan, A. B., Hamsher, K., Varney, N. R., & Spreen, O. (1994). *Contributions to Neuropsychological Assessment.* New York: Oxford University Press, (Chapter Chapter).

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews. Neuroscience, 3*, 243-249.

Bruce, V., & Young, A. W. (2012). *Face Perception*: Psychology Press, (Chapter Chapter).

Calder, A. J., Keane, J., Young, A. W., Lawrence, A. D., Mason, S., & Barker, R. A. (2010). The relation between anger and different forms of disgust: implications for emotion recognition impairments in Huntington's disease. *Neuropsychologia, 48*, 2719-2729.

Castle, E., Eisenberger, N. I., Seeman, T. E., Moons, W. G., Boggero, I. A., Grinblatt, M. S., & Taylor, S. E. (2012). Neural and behavioral bases of age differences in perception of trust. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 20848-20852.

Chiao, J. Y., Adams, R. B., Tse, P. U., Lowenthal, L., Richeson, J. A., & Ambady, N. (2008). Knowing Who's Boss: fMRI and ERP Investigations of Social Dominance Perception. *Group Process Intergroup Relat, 11*, 201-214.

Duff, K., Paulsen, J. S., Beglinger, L. J., Langbehn, D. R., & Stout, J. C. (2007). Psychiatric symptoms in Huntington's disease before diagnosis: the predict-HD study. *Biological Psychiatry, 62*, 1341-1346.

Dumas, E. M., van den Bogaard, S. J., Middelkoop, H. A., & Roos, R. A. (2013). A review of cognition in Huntington's disease. *Front Biosci (Schol Ed), 5*, 1-18.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, California: Consulting Psychologists Press, (Chapter Chapter).

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience, 19*, 1508-1519.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends Cogn Sci, 11*, 77-83.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15*, 870-878.

Gray, J. M., Young, A. W., Barker, W. A., Curtis, A., & Gibson, D. (1997). Impaired recognition of disgust in Huntington's disease gene carriers. *Brain, 120 (Pt 11)*, 2029-2038.

Hall, J., Whalley, H. C., McKirdy, J. W., Sprengelmeyer, R., Santos, I. M., Donaldson, D. I., McGonigle, D. J., Young, A. W., McIntosh, A. M., Johnstone, E. C., & Lawrie, S. M. (2010).

A common neural system mediating two different forms of social judgement. *Psychological Medicine, 40*, 1183-1192.

Hayes, C. J., Stevenson, R. J., & Coltheart, M. (2007). Disgust and Huntington's disease. *Neuropsychologia, 45*, 1135-1151.

Hennenlotter, A., Schroeder, U., Erhard, P., Haslinger, B., Stahl, R., Weindl, A., von Einsiedel, H. G., Lange, K. W., & Ceballos-Baumann, A. O. (2004). Neural correlates associated with impaired disgust processing in pre-symptomatic Huntington's disease. *Brain, 127*, 1446-1453.

Hoffman, P., Meteyard, L., & Patterson, K. (2014). Broadly speaking: Vocabulary in semantic dementia shifts towards general, semantically diverse words. *Cortex, 55*, 30-42.

Huntington-Study-Group. (1996). Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group. *Movement Disorders, 11*, 136-142.

Johnson, S. A., Stout, J. C., Solomon, A. C., Langbehn, D. R., Aylward, E. H., Cruce, C. B., Ross, C. A., Nance, M., Kayson, E., Julian-Baros, E., Hayden, M. R., Kieburtz, K., Guttman, M., Oakes, D., Shoulson, I., Beglinger, L., Duff, K., Penziner, E., & Paulsen, J. S. (2007). Beyond disgust: impaired recognition of negative emotions prior to diagnosis in Huntington's disease. *Brain, 130*, 1732-1744.

Keating, C. E. (1985). Human dominance signals: The primate in us. In S. L. Ellyson & J. E. Dovidio (Eds.), *Power, dominance, and nonverbal behavior* (Vol. Keating, C. E (1985). Human dominance signals: The primate in us. In S. L. Ellyson & J. E Dovidio (Eds.), Power, dominance, and nonverbal behavior (pp. 89-108). New York: Springer-Verlag., pp. 89-108). New York: Springer-Verlag.

Kunimatsu, A., Aoki, S., Masutani, Y., Abe, O., Hayashi, N., Mori, H., Masumoto, T., & Ohtomo, K. (2004). The optimal trackability threshold of fractional anisotropy for diffusion tensor tractography of the corticospinal tract. *Magnetic Resonance in Medical Sciences, 3*, 11-17.

McColgan, P., Seunarine, K. K., Razi, A., Cole, J. H., Gregory, S., Durr, A., Roos, R. A., Stout, J. C., Landwehrmeyer, B., Scahill, R. I., Clark, C. A., Rees, G., & Tabrizi, S. J. (2015). Selective vulnerability of Rich Club brain regions is an organizational principle of structural connectivity loss in Huntington's disease. Brain, 138, 3327-3344.

Mattavelli, G., Andrews, T. J., Asghar, A. U., Towler, J. R., & Young, A. W. (2012). Response of face-selective brain regions to trustworthiness and gender of faces. *Neuropsychologia, 50*, 2205-2211.

Milders, M., Crawford, J. R., Lamb, A., & Simpson, S. A. (2003). Differential deficits in expression recognition in gene-carriers and patients with Huntington's disease. *Neuropsychologia, 41*, 1484-1492.

Mitchell, I. J., Heims, H., Neville, E. A., & Rickards, H. (2005). Huntington's disease patients show impaired perception of disgust in the gustatory and olfactory modalities. *Journal of Neuropsychiatry and Clinical Neurosciences, 17*, 119-121.

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior, 27*, 237-254.

Montepare, J. M., & Dobish, H. (2014). Younger and older adults' beliefs about the experience and expression of emotions across the life span. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 69*, 892-896.

Mueller, H. P., Suessmuth, S. D., Landwehrmeyer, G. B., Ludolph, A. C., Tabrizi, S. J., Kloppel, S., & Kassubek, J. (2011). Stability effects on results of diffusion tensor imaging analysis by reduction of the number of gradient directions due to motion artifacts: an application to presymptomatic Huntington's disease. *PLoS Currents: Huntington Disease, B;3*, RRN 1292.

Mueller, H. P., Unrath, A., Sperfeld, A. D., Ludolph, A. C., Riecker, A., & Kassubek, J. (2007). Diffusion tensor imaging and tractwise fractional anisotropy statistics: quantitative analysis in white matter pathology. *Biomedical Engineering Online, 6*, 42.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*, 95-102.

Oldmeadow, J. A., Sutherland, C. A. M., & Young, A. W. (2013). Facial stereotype visualization through image averaging. *Social Psychological and Personality Science*, 615-623.

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends Cogn Sci, 18*, 566-570.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 11087-11092.

Papoutsi, M., Labuschagne, I., Tabrizi, S. J., & Stout, J. C. (2014). The cognitive burden in Huntington's disease: pathology, phenotype, and mechanisms of compensation. *Movement Disorders, 29*, 673-683.

Patterson, K., & Hodges, J. R. (1992). Deterioration of word meaning: implications for reading. *Neuropsychologia, 30*, 1025-1040.

Paulsen, J. S. (2011). Cognitive impairment in Huntington disease: diagnosis and treatment. *Current Neurology and Neuroscience Reports, 11*, 474-483.

Penney, J. B., Jr., Vonsattel, J. P., MacDonald, M. E., Gusella, J. F., & Myers, R. H. (1997). CAG repeat number governs the development rate of pathology in Huntington's disease. *Annals of Neurology, 41*, 689-692.

Poudel, G. R., Stout, J. C., Dominguez, D. J., Salmon, L., Churchyard, A., Chua, P., Georgiou-Karistianis, N., & Egan, G. F. (2014). White matter connectivity reflects clinical and cognitive status in Huntington's disease. Neurobiology of Disease, 65, 180-187.

Rees, E. M., Farmer, R., Cole, J. H., Henley, S. M., Sprengelmeyer, R., Frost, C., Scahill, R. I., Hobbs, N. Z., & Tabrizi, S. J. (2014). Inconsistent emotion recognition deficits across stimulus modalities in Huntingtons disease. *Neuropsychologia, 64C*, 99-104.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.

Said, C. P., Haxby, J. V., & Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366*, 1660-1670.

Santos, I. M., & Young, A. W. (2005). Exploring the perception of social characteristics in faces using the isolation effect. *Visual Cognition, 12*, 213-247.

Santos, I. M., & Young, A. W. (2008). Effects of inversion and negation on social inferences from faces. *Perception, 37*, 1061-1078.

Santos, I. M., & Young, A. W. (2011). Inferring social attributes from different face regions: evidence for holistic processing. *Quarterly Journal of Experimental Psychology, 64*, 751-766.

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. Taguiri & L. Oetrullo (Eds.), *Person perception and interpersonal behaviour* (pp. 300-315). Stanford, CA: Stanford University Press.

Sprengelmeyer, R., Orth, M., Muller, H. P., Wolf, R. C., Gron, G., Depping, M. S., Kassubek, J., Justo, D., Rees, E. M., Haider, S., Cole, J. H., Hobbs, N. Z., Roos, R. A., Durr, A., Tabrizi, S.

J., Sussmuth, S. D., & Landwehrmeyer, G. B. (2014). The neuroanatomy of subthreshold depressive symptoms in Huntington's disease: a combined diffusion tensor imaging (DTI) and voxel-based morphometry (VBM) study. *Psychological Medicine, 44*, 1867-1878.

Sprengelmeyer, R., Schroeder, U., Young, A. W., & Epplen, J. T. (2006). Disgust in pre-clinical Huntington's disease: a longitudinal study. *Neuropsychologia, 44*, 518-533.

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., Perrett, D. I., & Rowland, D. (1996). Loss of disgust. Perception of faces and emotions in Huntington's disease. *Brain, 119 (Pt 5)*, 1647-1665.

Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: ambient images generate a three-dimensional model. *Cognition, 127*, 105-118.

Sutherland, C. A. M., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychoiogy, 106*, 186-208.

Sutherland, C .A. M., Young, A. W., & Rhodes, G. (in press). Facial first impressions from another angle: how social judgments are influenced by changeable and invariant facial properties. *British Journal of Psychology*.

Tiddeman, B. P., Perrett, D. I., & Burt, D. M. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications, 21*, 42-50.

Todorov, A. (2008). Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences, 1124*, 208-224.

Todorov, A., & Engell, A. D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social Cognitive and Affective Neuroscience, 3*, 303-312.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*, 519-545.

Unrath, A., Muller, H. P., Riecker, A., Ludolph, A. C., Sperfeld, A. D., & Kassubek, J. (2010). Whole brain-based analysis of regional white matter tract alterations in rare motor neuron diseases by diffusion tensor imaging. *Human Brain Mapping, 31*, 1727-1740.

van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. Journal of Neuroscience, 31, 15775-15786.

Vernon, R. J., Sutherland, C. A., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences of the United States of America, 111*, E3353-3361.

Wang, K., Hoosain, R., Yang, R. M., Meng, Y., & Wang, C. Q. (2003). Impairment of recognition of disgust in Chinese with Huntington's or Wilson's disease. *Neuropsychologia, 41*, 527-537.

Willis, M. L., Palermo, R., Burke, D., McGrillen, K., & Miller, L. (2010). Orbitofrontal cortex lesions result in abnormal social judgements to emotional faces. *Neuropsychologia, 48*, 2182-2187.

Willis, M. L., Palermo, R., & Burke, D. (2011). Social judgments are influenced by both facial expression and direction of eye gaze. *Social Cognition, 29*, 415–429.

Willis, M. L., Palermo, R., McGrillen, K., & Miller, L. (2014). The nature of facial expression recognition deficits following orbitofrontal cortex damage. *Neuropsychology, 28*, 613-623.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5*, 277-283.

Young, A. W., Perrett, D. I., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial Expressions of Emotion: Stimuli and Tests (FEEST)*. Bury St. Edmunds, Suffolk: Thames Valley Test Company, (Chapter Chapter).

Zebrowitz, L. A., Kikuchi, M., & Fellous, J. M. (2007). Are effects of emotion expression on trait impressions mediated by babyfaceness? Evidence from connectionist modeling. *Personality and Social Psychology Bulletin, 33*, 648-662.