Facial cues to depressive symptoms and their associated personality attributions

Naomi Jane Scott *, Robin Stewart Samuel Kramer, Alex Lee Jones, Robert Ward

Bangor University, School of Psychology, Adeilad Brigantia, Gwynedd, LL57 2AS, UK

ABSTRACT

Depression is a common mental health disorder, with 12% of the UK population diagnosed at any one time. We assessed whether there are cues to depressive symptoms within the static, non-expressive face, and if other socially relevant impressions might be made by these cues. Composite "average" face images were created from students scoring high and low on self-report measures of depressive symptoms, capturing potential correlations between facial appearance and symptoms of depression. These were then used in a warping procedure, creating two versions of individual faces, one warped towards the high symptom composite, and the other towards the low. In Experiment 1, we first found observers were able to identify images representing high and low symptom occurrence at levels significantly greater than chance. Secondly, we collected observer impressions of the two versions of each face. The faces reflecting high levels of depressive symptoms were picked as less socially desirable over a broad range of personality trait estimates compared to low symptom images. In Experiment 2, we replicated the key finding that the static face contains cues to levels of depression symptoms, using composites created from a new database of student photos and depression inventory scores.

© 2013 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Depression is a disabling mental disorder which affects 12% of the adult UK population at any one time (Singleton et al., 2001). The effect of this widespread diagnosis is far-reaching; by 2030, depression is proposed to be the second largest cause of worldwide disability, rising to the leading cause in high-income countries (Mathers and Loncar, 2006). At an individual level, those with depression are more likely to have fewer qualifications, a lower socioeconomic status (Singleton et al., 2001), and form a group that have a higher suicide incidence rate than the general population (Simon and Vonkorff, 1998). Depression also has a negative impact on the economy, with patients requiring more doctors, community care and medication than others (Singleton et al., 2001). A deeper understanding of the ability to discriminate between individuals with and without depressive symptoms, alongside how this affects perceptions of social desirability, may aid knowledge of the social implications of perceived depression.

There are two main types of cues that could be utilised to help discriminate between depressed and non-depressed individuals. The first considered here is that of live cues, where participants are able to use cues from body motions. In a study of "thin slices" of non-verbal behaviours and appearance, Waxer (1976) showed two-minute silent video segments of depressed and non-depressed patients to undergraduate and clinical graduate students. The patients' Minnesota Multiphasic Personality Inventory depression subscale scores and participants' depression ratings showed a correlation of 0.60, suggesting an ability to recognise depressive symptoms from non-verbal cues. These live cues included information from physical movements such as body motion, facial expressions, posture and proximity to others which could help to identify individuals with depression. Use of these live cues has been explored in other studies; for example, Michalak et al. (2009) found that body motion in minimal point-light displays was a reliable cue to depression, with depressed participants walking more slowly and with less amplitude in limb movements. Geerts and Brüne (2011) discuss findings of other live cues of depression, such as the ability to predict depression course from the impaired social behaviour of depression-prone patients.

However, recent research indicates an unexpected richness from a second type of cue — static cues. An example is the static, non-expressive face. Many of the Big Five personality traits, as measured by self-report, can be accurately estimated from "passport"-type photographs of people looking straight-ahead and with a neutral expression, both using individual photos (Penton-Voak et al., 2006) and composite images comparing the highest and lowest trait levels (Little and Perrett, 2007; Kramer and Ward, 2010). Composite images are computer-generated averaged faces of a number of individuals. These composite images reduce

individual variation and preserve the common facial features of those within the group. Besides Big Five personality traits, the face can also reliably cue socially relevant traits such as socio-sexual orientation (Boothroyd et al., 2008) and trustworthiness (Stirrat and Perrett, 2010).

Limited research has been conducted into the ability to recognise clinically related conditions from facial images. Shevlin et al. (2003) demonstrated that participants were able to accurately judge levels of psychoticism from facial photographs. However, as jewellery, clothing and hairstyles of the stimuli were not kept consistent, it cannot be determined whether the accuracy reported is a product of facial cues, or aided by additional cues such as clothing type. Research by Holtzman (2011) has helped to overcome these criticisms by using images in which individuals have removed all make-up and jewellery, and tied their hair away from their face. Even after removing these confounds, participants were able to accurately identify the 'dark triad' of personality disorders (machiavellians, narcissists and psychopaths) from composite facial images.

Results like those above raise the question of whether reliable cues for symptoms of depression (and other mental health conditions) may be present in the non-expressive face. The possibility of a link between facial appearance and symptoms of depression is strengthened by two findings. First, high trait Neuroticism is a risk factor for depression; and second, trait Neuroticism is one of the factors reliably signalled on the face. We discuss these findings in turn.

Personality traits are heritable constructs that serve as risk factors for mental health disorders (Krueger et al., 2000; Sen et al., 2003). Neuroticism has been identified as a personality trait with special importance for depression, as it may increase the likelihood of an individual experiencing distress (Bagby et al., 1996), which can in turn manifest itself as anger, anxiety or depression (Bagby et al., 1997). Indeed, individuals with depression are consistently found to have higher than average levels of Neuroticism (Widiger and Trull, 1992; Trull and Sher, 1994; Bagby et al., 1997). Although there is controversy surrounding the direction of the neuroticism–depression link, the general consensus suggests that Neuroticism forms a diathesis for depression (Trull and Sher, 1994; Krueger et al., 2000). The ability to use Neuroticism levels to predict later depression provides some supporting evidence for this theory (Roberts and Kendler, 1999).

Neuroticism also seems to be one of the traits which is most reliably signalled from the neutral face. Kramer and Ward (2010) found that composite images of women with high and low levels of trait Neuroticism could be identified at levels significantly greater than chance, even from internal facial features alone. In fact, Kramer and Ward (2010) found that removing the outer facial features (top of the head, jaw-line, etc.), leaving only the eyes, nose, and mouth areas within the images, did not impair identification of Neuroticism. More recently, Jones et al. (2012) created high and low Neuroticism composite images on the basis of 3D facial scans, to remove potential postural or alignment cues from subsequent comparisons. Again, identification of Neuroticism composites was robust even without these cues.

We sought to determine whether symptoms of depression might be similarly cued in the static, non-expressive face in two experiments. In Experiment 1, we tested the possibility that, within a non-clinical sample, people who score high for symptoms of depression have a distinct facial appearance from those who score low for depressive symptoms, and if so, whether such differences in appearance can be identified by untrained observers. We also explored how appearance impacted observer impressions of "social desirability", that is, whether the person depicted in the photo was perceived to be a desirable social partner. In particular, we examined whether the facial appearance

correlated with depressive symptoms might leave negative impressions on observers. These observer impressions might be important even if they are not accurate, as they could mean that people at risk of depression receive unfavourable social reactions, even in the absence of overt depressive behaviour. This unfavourable reaction could then increase social isolation and negative social consequences associated with depression. In Experiment 2, we created a new database of photographs and depression inventory scores, to replicate the identification task used in Experiment 1, with a new set of composite images.

2. Experiment 1

2.1. Method

This experiment consisted of two phases, the first of which ("stimulus creation") involved the collection of photographs and a measure of depressive symptoms to create our face stimuli. In the second phase, these stimuli were used to investigate the ability of observers to discriminate between faces representing high and low depression symptom occurrence (Experiment 1a), and to measure observer impressions from the high and low images of important social traits (Experiment 1b). All stages of the study were approved by Bangor University's departmental ethics committee.

2.1.1. Stimulus creation

Our aim in this phase was to create stimuli which reflected the actual differences, if any, in the faces of people who report high and low levels of depressive symptoms. To do so, we created composite ("averaged") images from the faces of people who reported they had high levels of depressive symptoms, and another set of composites from the faces of people reporting low levels of symptoms. These composites were then used as anchor points to let us create new images, by warping individual faces towards the high and low depression appearances.

2.1.1.1. Participants. Two hundred and twenty five Bangor University students (130 females, age $M=21.45$, S.D.$=5.04$) were recruited and paid £5 for their participation.

2.1.1.2. Measures. Depressive symptoms were assessed using questions based on the revised Beck's depression inventory (BDI, Beck et al., 1996), in which depression symptom severity was indicated on a 4-point Likert scale ranging from 0 (disagree) to 3 (very much agree). Summed scores on the BDI can give an indication of depressive symptom severity, with scores between 0 and 13 indicating minimal depression, 14 and 19 mild, 20 and 28 moderate and 29 and 63 severe depression (Beck et al., 1996). Question nine, asking participants about suicidal thoughts, was removed in accordance with the Bangor University's departmental ethics committee, resulting in a 20 item questionnaire with a possible score range of 0–60. This adaption did not appear to impact upon the reliability of the questionnaire (Cronbach's $\alpha=0.91$). Male scores ranged from 0 to 49, $M=14.94$, S.D.$=12.36$, female scores ranged from 0 to 43, $M=16.41$, S.D.$=10.53$. These scores were later used to separate individuals reporting the highest and lowest levels of depression symptoms; purely for simplicity of exposition, these will be referred to as depressed and non-depressed groups, although to be clear, these were both non-clinical samples.

2.1.1.3. Procedure. Facial photographs were taken with a professional camera from a distance of 2 m whilst camera height, zoom and flash were kept consistent. Participants were

asked to sit down in front of a clear, light coloured background and remove make-up and jewellery, tie their hair back and adopt a neutral facial expression with eyes facing the camera. Individuals whose self-reported ethnicity was not White British, those aged over 30, and males with facial hair were removed in order to maintain composite quality, resulting in a sample of 106 females and 48 males.

The fifteen highest (female score $M=34.73$, S.D.$=5.60$, male $M=27.73$, S.D.$=9.07$) and lowest scorers (female score $M=2.53$, S.D.$=1.46$, male $M=3.13$, S.D.$=2.23$) on the depression survey were then selected for each sex, and used to produce high and low depression composites for each sex (Fig. 1) using JPsychoMorph software (Tiddeman et al., 2001).

The composite faces shown in Fig. 1 capture statistical regularities in the facial appearance of people who reported high and low levels of depressive symptoms. These composites were not shown to observers, but were used in the next part of the stimulus creation process, as anchor points to create sexually-specific gradients of shape and texture changes related to high and low depression. Forty individuals (20 males, 20 females) with a depression score close to the group average, and who had given consent to have their photograph used individually, were picked at random from the photograph dataset. Each individual's image was warped 50% towards the high and also 50% towards the low

depression composites (Rowland and Perrett, 1995). This resulted in two images for each individual, one transformed to reflect the statistics of face shape and texture in the high depressed composite, the other the low (see Fig. 2 for examples). The 40 stimulus pairs created were used in the following discrimination experiments.

2.1.2. Experiment 1a: Can depressive symptoms be identified from the static face?

In this experiment, we assessed observers' ability to accurately discriminate between the 40 facial image pairs created in Phase 1. Within each pair, a facial photo was transformed to produce one image reflecting the statistics of face shape and texture relating to high levels of depressive symptoms; the other, low levels.

2.1.2.1. Method. Observers:
Sixty-four undergraduate Bangor University students aged 18–41 (46 females, age $M=19.90$, S.D.$=4.03$) participated in the experiment for course credit.

Procedure: Observers were presented with each warped image pair (38×50 mm on screen) side by side (position counter-balanced) on a computer screen below a discriminative statement. Observers were asked to select which image had 'more symptoms of depression' for all image pairs. Responses were not

Fig. 1. (Colour online) Composite images created from the photographs of the highest and lowest 15 male and female depression scorers.

High depression   **Low depression**

Fig. 2. (Colour online) Warped image pair examples. Images warped 50% towards the high and low depression composite images illustrated in Fig. 1.

speeded, and viewing distance was not fixed. After selecting an image to match the discrimination statement using a mouse click, the next trial was presented. The 40 trials were presented in a randomised order for each observer.

Statistical analysis: A 2 (sex of observer) × 2 (sex of stimuli) mixed ANOVA was conducted to determine if the sex of observer, sex of stimuli or an interaction between the two had significant impact on discrimination accuracy. Following this, a one-sample t-test was used to compare observer accuracy to chance performance. Chance performance in this two-alternative, forced-choice discrimination task was 50%.

2.1.2.2. Results. Correct responses were counted when observers selected the high depression image as having more symptoms of depression. A 2 × 2 mixed ANOVA showed no significant effect of stimulus sex, $F(1, 62)=0.74$, $p=0.40$, observer sex, $F(1, 62)=0.18$, $p=0.67$, or interaction effects, $F(1, 62)=0.30$, $p=0.86$, suggesting that there was no difference in accuracy between male and female observers, or between male and female stimuli. For these reasons results are presented aggregated over stimulus and observer sex. Observers were able to accurately discriminate high and low depression images 61% of the time, a level significantly greater than chance, $t(63)=9.81$, $p < 0.000001$.

2.1.3. Experiment 1b: Perceptions of social traits in faces high and low for depressive symptoms

Experiment 1a showed that observers were able to discriminate between images reflecting the faces of people with high and low levels of depressive symptoms more accurately than expected by chance. Given that untrained observers were sensitive to the correlation between facial appearance and depressive symptoms, it is interesting to ask what other perceptions these faces might afford. In Experiment 1b we explored the impressions of socially-relevant traits attributed to faces representing high and low depression. That is, what kinds of social impressions would observers receive from a face reflecting the statistical properties of high or low levels of depressive symptoms? For example, would observers have the impression that depressed faces were less emotionally stable or agreeable? These impressions made by the face are relevant as they may bias the social responses of observers. We assessed perceptions of Big Five personality traits, as a summary description of socially-relevant impressions. Attractiveness and perceived gender were also included as important socially relevant variables that are driven by appearance.

2.1.3.1. Method. Observers: A new sample of 48 Bangor University undergraduate students aged 18–33 (37 females, age $M=19.78$, S.D.$=2.39$) took part in the experiment for course credit.

Procedure: The 40 pairs of high and low depression images were presented in a similar way to Experiment 1. In this case however, each pair was presented with seven different discrimination statements, for a total of 280 trials. The seven discrimination statements consisted of one statement for each of the Big Five personality factors, one for attractiveness, and one for perceived femininity. For most cases, apart from gender and Neuroticism, observers were asked to select the face which had the more socially desirable appearance: more Extraverted ("Who looks more outgoing and energetic?"), Agreeable ("more friendly and compassionate"), Conscientious ("Who looks more efficient and organised?"), Open ("Who looks more open to new experiences?"), and attractive ("Who is more attractive?"). For the Neuroticism discrimination, observers were asked "Who looks more sensitive and nervous?" and for the gender discrimination observers were asked to select the face which was more feminine. This final discrimination statement was chosen to explore the perceived gender of the stimuli, assuming masculinity and femininity are at opposite ends of a continuous spectrum.

To reduce chances of fatigue, the experiment was separated into two blocks, with a short break offered in-between. One block contained the female images and the other the males, presented in a counterbalanced order across observers. The blocks were split in this way as the male and female stimuli were created separately, based on different sets of anchor points. Within a block, face pairs and discrimination statements appeared in randomised order for each observer. Between blocks, a break was given as needed.

Statistical analysis: A 7(Discrimination statement) $\times$ 2(Sex of stimuli) $\times$ 2(Sex of observer) factorial MANOVA (robust to violations of sphericity) was used to determine the effect of observer and stimuli sex on observations. Following this, one-sample *t*-tests were used to compare observers accuracy to 50% chance level. Due to the use of multiple t-tests a Holm's sequential Bonferroni correction (Holm, 1979) was applied, resulting in a stricter alpha level and reducing the chance of type-one errors. Additionally correlation analysis was used to determine the possible influence of perceived attractiveness; again a Holm's sequential Bonferroni correction was used to compensate for multiple tests.

2.1.3.2. Results and discussion.
For easy comparison, scores for Neuroticism were reversed to represent Emotional Stability, so that all personality traits now increase in desirability as scores increase.

A 7 $\times$ 2 $\times$ 2 factorial MANOVA showed only the expected main effect of trait, $F(6, 41)=4.28$, $p=0.002$, but no significant effect of stimulus sex, $F(1, 46)=1.67$, $p=0.69$, observer sex, $F(1, 46)=0.18$, $p=0.67$, or interaction between the two, $F(1, 46)=0.28$, $p=0.60$, or any other interaction, all F's$(6, 41)<1.80$, $p>0.05$. We therefore carried out subsequent analyses separately for each trait, aggregating over stimulus sex and observer sex.

As illustrated in Fig. 3, when presented with high and low symptom versions of the same face, observers attributed more positive personality traits to the facial images of low depression, at a rate significantly higher than expected by chance. Compared to high depression images, low depression images were selected as more Agreeable, $t(47)=5.58$, $p=0.00000069$, Conscientious, $t(47)=2.41$, $p=0.0199$, Emotionally stable, $t(47)=5.07$, $p=0.0000034$, Extraverted, $t(47)=4.26$, $p=0.00039$, and Open, $t(47)=7.26$, $p=0.000000023$, corrected for multiple comparisons. These results suggest that observers may form a negative impression from the facial appearance associated with depression.

When compared to high depression images, observers also perceived the low depression images to be both more attractive and feminine 55% of the time, again a rate significantly greater than expected by chance, $t(47)=3.30$, $p=0.0054$ and $t(47)=3.24$,
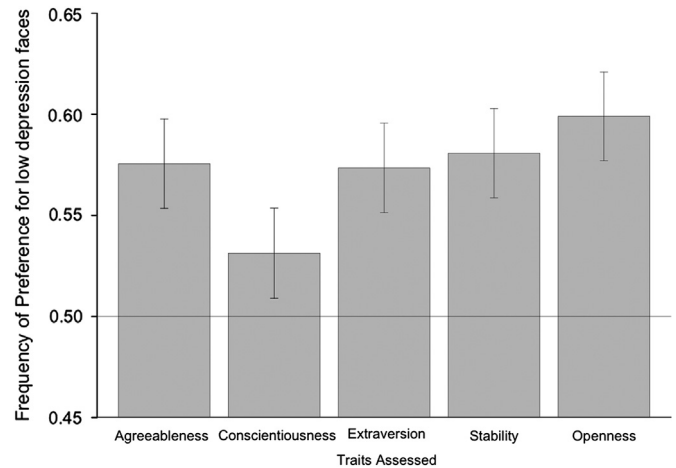

Fig. 3. Frequency of preference for low compared to high depression faces. In Experiment 1b, observers were presented with a high and low depression face, and asked which appeared more desirable for the traits assessed. Chance selection would therefore result in 50% selection. Error bars represent the 95% confidence interval of the mean for each trait.

$p=0.0044$ (corrected for multiple comparisons) respectively. Attractive individuals are often perceived to be high in desirable traits, a concept known as the halo effect (Dion et al., 1972). As low symptom occurrence faces were perceived to be more attractive, as well as higher in desirable personality traits, we conducted a series of correlation analyses to determine if the halo effect was the driving force behind our findings. Although correlations were generally positive, only perceptions of Openness and attractiveness were found to be significantly correlated, $r=0.62$, $p=0.00002$, all other correlations were insignificant, all r's <0.34, all p's >0.08, corrected for multiple comparisons. Despite the association between perceived levels of Openness and attractiveness, the absence of significant correlations between the other personality traits and attractiveness suggests that an attractiveness halo is unlikely to be the driving force behind the wide-spread attribution of more negative traits to the high depression images.

3. Experiment 2

3.1. Identifying levels of depressive symptoms from the static face in a new sample

In Experiment 1 we demonstrated that observers were able to identify, at above-chance levels, facial images representing high and low symptoms of depression. We thought it was important to replicate and generalise this finding. We sought to replicate these results by creating a new database of face images, and generating new high and low composite and warped images. We also wanted to be certain that our results were not sensitive to the depression inventory used, and so this time levels of depressive symptoms were measured using the Inventory of Depressive Symptoms (IDS, Rush et al., 1996).

3.1.1. Method
This experiment was run using the same procedures as Experiment 1a, with the differences noted below.

3.1.1.1. Stimulus creation.
Stimulus creation proceeded exactly as in Experiment 1, with the following modifications. First, a new set of 221 Bangor University students (130 females, age $M=21.65$, S.D.$=5.09$) were recruited and paid £6 for their participation.

Second, the Inventory of Depressive Symptoms (IDS, Rush et al., 1996) was used as the measure of depressive symptoms. The IDS is a 30 item, self-rated, questionnaire where participants are asked to describe symptom severity on a 4 point scale, ranging from 0 suggesting no symptom presence, to 3 suggesting high symptom presence. Possible scores range from 0 to 84, with scores 0–25 suggesting no depression, 26–38 mild depression, 39–48 moderate depression and 49–84 severe depression. Female scores ranged from 1 to 62 ($M=17.48$, SD$=11.56$), male scores ranged from 1 to 43 ($M=15.41$, S.D.$=9.03$). The images of the 15 highest (female score $M=30.40$, S.D.$=8.40$, male $M=25.40$, S.D.$=7.69$) and 15 lowest (female score $M=5.40$, S.D.$=1.74$, male $M=6.4$, S.D.$=2.36$) scorers on the IDS were selected and then used to create composite stimuli for each sex. Following this, 20 male and 20 female images were selected to create the warped images, using the same procedure as in Experiment 1.

3.1.1.2. Observers. Forty-eight (36 females) undergraduate Bangor University students took part in the study for course credits. Observers' ages ranged from 18 to 36 ($M=20.54$, S.D.$=3.05$).

3.1.1.3. Procedure. The experimental procedure was identical to Experiment 1a.

3.1.2. Results and discussion

Answers were coded as correct when observers selected the image representing high depressive symptoms. A 2 (Sex of stimuli) × 2 (Sex of observer) mixed ANOVA showed no significant effect of observer sex, $F(1,46)=1.05$, $p=0.31$, stimulus sex, $F(1,46)=0.45$, $p=0.50$, or interaction, $F(1,46)=0.04$, $p=0.84$. For this reason, means were aggregated over stimulus and observer sex. Observers were accurate in their discriminations 58% of the time, a rate significantly more accurate than expected by chance $t(47)=3.86$, $p=0.0003$. An independent samples t-test found no difference in accuracy between Experiments 1a and 2, $t(110)=1.68$, $p=0.10$. Observers were therefore consistently able to identify facial images of individuals scoring high and low on measures of depression, significantly more often than expected by chance. These findings replicated over two independent stimulus samples, different depression measures, and different observer samples.

4. General discussion

We created images which captured statistical regularities in the facial appearance of students who did and did not report high levels of depressive symptoms - less formally, depressed and non-depressed faces. We found in two separate studies that untrained observers were able to discriminate between these images more accurately than expected by chance. These results demonstrate two important conclusions about facial appearance and depression. First, there are differences in the facial appearance of people based on their reported levels of depressive symptoms. Second, observers are sensitive to these differences: not only were observers able to discriminate between the high and low depression images more accurately than expected by chance, but they also made more negative social attributions towards the high depression images.

One range of possibilities is that there are biological factors, for example, genetic or hormonal, which influence both susceptibility to depression and facial appearance. Additional possibilities are related to a "self-fulfilling prophecy", in which people with a particular appearance learn through interaction with others to behave in a particular way. Here we briefly consider both possibilities.

The biological basis of depression is under active study, yet heritability estimates of Neuroticism are substantial, in the range

of 40% (Lake et al., 2000) to 50% (Floderus-Myrhed et al., 1980). Interestingly depression has also been shown to have a genetic aspect, with an estimated heritability of 31–42%, and additional findings that first-degree relatives of patients are more likely to have recurrent depression (Sullivan et al., 2000). It is possible that this is a reflection of the genetic transmission of Neuroticism that puts individuals at increased risk of depression development (Bagby et al., 1997). The type of depression discussed here would be seen as trait depression, where an individual has a stable trait (such as Neuroticism) which increases the likelihood of experiencing severe depression. However, depression can also occur as a state - where symptoms fluctuate over time (Teasdale, 1998). It is often hard to tease apart these disorders, however recent research conducted by Endler et al. (2000) suggests that measures such as the Beck's Depression Inventory (Beck et al., 1996) are accurate in measuring trait levels of depression. This means that the stimuli developed in this study are more likely to represent those who experience trait, as opposed to state, depression.

It is also clear that the face is a sensitive marker for many kinds of congenital and acquired disorders. Facial appearance is a diagnostic criterion in many such disorders (e.g., Downs and foetal alcohol syndromes). Williams syndrome is an example of a genetic disorder which presents a behavioural pattern of hyper-sociality alongside distinct morphological facial features (an 'elfin'shaped face). Although such results do not speak to our specific case of depression and the face, they do demonstrate that behavioural traits and facial morphology can be correlated in a highly specific manner. In short, a genetic or hormonal link between depressive symptoms and facial appearance seems like a plausible story, but there is little direct evidence for it at this point.

Our results bear more directly on the possibility of a cultural or acquired link between appearance and depression. Personality traits are important factors in determining friendship and social inclusion. Friends of highly extraverted individuals rated their friendship as closer, while friends of highly agreeable individuals reported less irritation and friendships with highly open individuals had fewer reported conflicts (Berry et al., 2000). Furthermore, individuals scoring highly on Agreeableness had higher numbers of new acquaintances wanting to form friendships with them (Selfhout et al., 2010). This literature suggests that individuals high in these personality traits (Agreeableness, Extraversion and Openness) are seen as more desirable friends, traits which observers perceived to be less prevalent in individuals with high symptom occurrence. Given these facts, it is highly relevant that we find the social response to the faces of people with high levels of depressive symptoms to be generally negative: for example observers perceived images of high depressive symptoms to be less Agreeable, less Conscientious, less Extraverted, less Open to experience, and less Emotionally Stable. That is, the high depression faces seem to be involuntarily broadcasting a message of negative personality traits associated with social undesirability, even in the absence of any overt behaviour. At the very least, this would mean that people with a prototypical high-depression face would face barriers and challenges for social inclusion that others would not. A failure to meet these social challenges could heighten any risk of depression, by making social exclusion more likely. This means that an individual's depressive symptoms may cause them to feel like they are being excluded socially, whilst the facial cues picked up by others casts a negative social light on them. Indeed findings from Segrin and Dillard (1992) that mild symptoms of depression can induce rejection from others further support this argument.

Our results demonstrate that facial appearance can hold cues for symptoms of depression. Additionally we have shown that observers perceive the facial appearance associated with depressive

symptoms to be less socially desirable. There are limitations to our study which are important to highlight. The photograph databases are collected from a non-clinical student sample, so it will be important to investigate the association between facial appearance and high levels of depression symptoms in a clinical population. This would help to determine the extent to which facial appearance is a useful cue for identification of depression. Secondly, although discriminations between high and low depressive symptom images are more accurate than expected by chance, performance is far from perfect. However, there are evidently cues within the face correlated to levels of depressive symptoms which further knowledge about the correlation between appearance and mental health, and what the visible cues may be that underlie this correlation.

In practical terms, given that observers formed negative impressions from the facial appearance statistically associated with depression, it may be that facial appearance variables are associated with social outcomes. For example, a face which approaches the high-depression average may be seen as having fewer socially desirable personality traits, which in turn could negatively affect social interactions.

Acknowledgement

References

Bagby, R.M., Bindseil, K.D., Schuller, D.R., Rector, N.A., Young, T., Cooke, R.G., Seeman, M.V., McCay, E.A., Joffe, R.T., 1997. Relationship between the five-factor model of personality and unipolar, bipolar and schizophrenic patients. Psychiatry Research 70 (2), 83–94.

Bagby, R.M., Young, L.T., Schuller, D.R., Bindscil, K.D., Cooke, R.G., Dickens, S.E., 1996. Bipolar disorder, unipolar depression and the five-factor model of personality. Journal of Affective Disorders 41 (1), 25–32.

Beck, A.T., Steer, R.A., Brown, G.K., 1996. BDI-II Manual. Psychological Corporation, San Antonia, TX.

Berry, D.S., Willingham, J.K., Thayer, C.A., 2000. Affect and personality as predictors of conflict and closeness in young adults' friendships. Journal of Research in Personality 34 (1), 84–107.

Boothroyd, L.G., Jones, B.C., Burt, D.M., DeDruine, L.M., Perrett, D.I., 2008. Facial correlates of sociosexuality. Evolution and Human Behavior 29 (3), 211–218.

Dion, K., Berscheid, E., Walster, E., 1972. What is beautiful is good. Journal of Personality and Social Psychology 24 (3), 285–290.

Endler, N.S., Macrodimitris, S.D., Kocovski, N.L., 2000. Depression: the complexity of self-report measures. Journal of Applied Biobehavioural Research 5 (1), 26–46.

Floderus-Myrhed, B., Pedersen, N., Rasmuson, I., 1980. Assessment of heritability for personality, based on a short-form of the Eysenck Personality Inventory: a study of 12,898 twin pairs. Behavior Genetics 10 (2), 153–162.

Geerts, E., Brüne, M., 2011. On the role of ethology in clinical psychiatry: what do ontogene and causal factors tell us about ultimate explanations of depression? In: Adriaens, P.R., De Block, A. (Eds.), Maladapting Minds: Philosophy, Psychiatry, and Evolutionary Theory. Oxford University Press, New York, pp. 117–142.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6 (2), 65–70.

Holtzman, N.S., 2011. Facing a psychopath: detecting the dark triad from emotionally-neutral faces, using prototypes from the personality faceaurus. Journal of Research in Personality 45 (6), 648–654.

Jones, A.L., Kramer, R.S.S., Ward, R., 2012. Signals of personality and health: the contributions of facial shape, skin texture, and viewing angle. Journal of Experimental Psychology: Human Perception and Performance 38 (6), 1353–1361.

Kramer, R.S.S., Ward, R., 2010. Internal facial features are signals of personality and health. The Quarterly Journal of Experimental Psychology 63 (11), 2273–2287.

Krueger, R.F., Caspi, A., Moffitt, T.E., 2000. Epidemiological personology: the unifying role of personality in population-based research on problem behaviours. Journal of Personality 68 (6), 967–998.

Lake, R.E., Eaves, L.J., Maes, H.H.M., Heath, A.C., Martin, N.G., 2000. Further evidence against the environmental transmission of individuals differences in neuroticism from a collaborative study of 45, 850 twins and relatives on two continents. Behavior Genetics 30 (3), 223–233.

Little, A.C., Perrett, D.I., 2007. Using composite images to assess accuracy in personality attribution to faces. British Journal of Psychology 98 (1), 111–126.

Mathers, C.D., Loncar, D., 2006. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Medicine 3 (11), 2011–2030.

Michalak, J., Troje, N.F., Fischer, J., Vollmar, P., Heidenreich, T., Schulte, D., 2009. Embodiment of sadness and depression: gait patterns associated with dysphoric mood. Psychomatic Medicine 71 (5), 580–587.

Penton-Voak, I.S., Pound, N., Little, A.C., Perrett, D.I., 2006. Personality judgments from natural and composite facial images: more evidence for a "kernel of truth" in social perception. Social Cognition 24 (5), 607–640.

Roberts, S.B., Kendler, K.S., 1999. Neuroticism and self-esteem as indices of the vulnerability to major depression in women. psychological Medicine 29 (5), 1101–1109.

Rowland, D.A., Perrett, D.I., 1995. Manipulating facial appearance through shape and colour. IEEE Computer Graphics and Applications 15 (5), 70–76.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., Trivedi, M.H., 1996. The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychological Medicine 26, 477–486.

Segrin, C., Dillard, J.P., 1992. The interactional theory of depression: a meta-analysis of the research literature. Journal of Social and Clinical Psychology 11 (1), 43–70.

Selfhout, M., Burk, W., Branje, S., Denissen, J., van Aken, M., Meeus, W., 2010. Emerging late adolescent friendship networks and Big Five personality traits: a social network approach. Journal of Personality 78 (2), 509–538.

Sen, S., Nesse, R.M., Stoltenberg, S.F., Li, S., Gleiberman, L., Chakravartim, A., Weder, A.B., Burmeister, M., 2003. A BDNF coding variant is associated with the NEO personality inventory domain neuroticism, a risk factor for depression. Neuropsychopharmacology 28, 397–401.

Shevlin, M., Walker, S., Davies, M.N.O., Banyard, P., Lewis, A.C., 2003. Can you judge a book by its cover? Evidence of self-stranger agreement on personality at zero acquaintance. Personality and Individual Differences 35 (6), 1373–1383.

Simon, G.E., VonKorff, M., 1998. Suicide mortality among patients treated for depression in an insured population. American Journal of Epidemiology 147 (2), 155–160.

Singleton, N., Bumpstead, R., O'Brien, M., Lee, A., Meltzer, H., 2001. Psychiatric morbidity Among Adults Living in Private Households, 2000. The Stationary Office, London.

Stirrat, M., Perrett, D.I., 2010. Valid facial cues to cooperation and trust: male facial width and trustworthiness. Psychological Science 21 (3), 349–354.

Sullivan, P.F., Neale, M.C., Kendler, K.S., 2000. Genetic epidemiology of major depression: review and meta-analysis. The American Journal of Psychiatry 157 (10), 1152–1562.

Teasdale, J.D., 1998. Cognitive vulnerability to persistent depression. Cognition and Emotion 2 (3), 247–274.

Tiddeman, B.P., Burt, D.M., Perrett, D.I., 2001. Prototyping and transforming facial textures for perception research. IEEE Computer Graphics and Applications 21, 42–45.

Trull, T.J., Sher, K.J., 1994. Relationship between the five-factor model of personality and axis I disorders in a nonclinical sample. Journal of Abnormal Psychology 103 (2), 350–360.

Waxer, P., 1976. Nonverbal cues for depth of depression: set versus no set. Journal of Consulting and Clinical Psychology 44 (3), 493.

Widiger, T.A., Trull, T.J., 1992. Personality and psychopathology: an application of the five-factor model. Journal of Personality 60 (2), 363–393.