I feel bad and look worse than you: Social comparisons moderate the effect of mood on face health judgement.

Laura Mirams[a], Ellen Poliakoff[a], Elizabeth H. Zandstra[b], Marco Hoeksma[b], Anna Thomas[c] and Wael El-Deredy[a].

[a]School of Psychological Sciences, Zochonis Building, University of Manchester
L.R.Mirams@ljmu.ac.uk,[1] ellen.poliakoff@manchester.ac.uk, wael.el-deredy@manchester.ac.uk.
[b]Unilever R&D Vlaardingen, The Netherlands.
liesbeth.zandstra@unilever.com, marco.hoeksma@unilever.com.
[c]Unilever R&D, Port Sunlight, UK
anna.thomas@unilever.com.

NO. TEXT PAGES: 31 (WORDS, 7,373, CHARACTERS, 40,007, 2 FIGURES

RESUBMITTED TO ACTA PSYCHOLOGICA AS AN ORIGINAL ARTICLE

CORRESPONDENCE TO: Dr Laura Mirams, School of Natural Sciences and Psychology, Room 2.23 Tom Reilly Building, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, U.K.
Tel: +44 (0) 151 904 6292
Email: L.R.Mirams@ljmu.ac.uk

[1] The research was conducted at the University of Manchester. The first author has since moved to Liverpool John Moores University.

ABSTRACT (179)

Mood can bias the judgements people make about themselves and how people compare themselves to others. However, it is not yet clear whether mood also affects appearance-based self-evaluations and social comparisons. Given the importance of perceived health status for well-being, we investigated the effect of mood on self-image and social comparisons of healthiness during two versions of a face health judgement task. Thirty participants judged how they felt compared to healthy and unhealthy looking versions of their own (self version) and a stranger's face (stranger version), after a positive, negative and neutral mood induction. Indeed, the effect of mood was dependent on self/stranger task order. Although mood did not affect face health judgement for participants who initially judged themselves against their own face, it did affect face health judgement for participants who initially judged themselves in comparison to a stranger's face. After the positive and negative mood inductions, these participants judged themselves as equivalent to healthier/unhealthier looking versions of their own and stranger's faces, respectively. Thus, social comparisons of facial healthiness could provide a perceptual measure of state well-being.

KEYWORDS: Mood, Self-evaluation, Social comparisons, Health perception, Well-being, Self-image

1. INTRODUCTION

Although we might think of ourselves as rational observers and decision makers, evidence suggests that the judgements that we make about ourselves, including our appearance, do not just depend on the visual information that comes in through our senses. Rather, higher level (or "top-down") factors can bias self-perceptions. For example, Epley and Whitchurch (2008) found that healthy people, particularly those with high self-esteem, perceive themselves to look more attractive than they do in reality. Mirams et al. (2014) also found that people can be biased in their perceptions of self-healthiness; compared to happy participants, those with dispositional tendencies to experience negative moods judged themselves as equivalent to less healthy looking versions of their own and another person's face. The purpose of the current study was to investigate whether experimentally induced, transient mood also moderates face health judgement and whether performing mood congruent social comparisons exacerbates the effect of mood on self-evaluations of healthiness.

Whereas the term emotion is often used to refer to intense, brief, differentiated states (e.g., fear, anger, joy, contentedness), moods are less intense, longer-lasting, and tend to be differentiated on valence and/or arousal (e.g., a good versus bad mood), rather than specific emotions (Sedikides, 1992; Winkielman, Knutson, Paulus, & Trujillo, 2007). The majority of evidence suggests that two dimensional models (e.g., Watson & Tellegen, 1985) best represent variations in emotional experience associated with mood states. According to Watson and Tellegen's two factor model, mood states can be represented along two independent dimensions of positive and negative affect (PA and NA). A state of high PA is characterised by pleasant feelings such as enthusiasm and alertness; with low PA associated with feelings of sadness and lethargy. A state of high NA, on the other hand, is characterised by feelings of psychological distress, such as nervousness and irritability, with low NA associated with feelings of calmness and serenity.

Mood states are thought to provide "a general, non-interruptive context for cognition" (Simon, 1982, p335), "gently colouring" and directing our ongoing thoughts and actions (Isen, 1984, p186-187). It has been argued that moods influence our judgements by priming mood congruent constructs in memory (e.g., Bower, 1981; M. S. Clark & Isen, 1982). According to these priming accounts, when we are feeling happy, positive thoughts about the self and others are more active and accessible in memory. When we are unhappy, negative thoughts are more active and accessible (Brown & Mankowski, 1993). Affect as Information theories (Clore, Schwarz, & Conway, 1994; Schwartz & Clore, 1988), on the other hand, suggest that we use current feeling states as information to inform our judgements; such that when we are feeling happy, we might misattribute positive affect to the target we are evaluating. These two explanations are considered to be complementary and have since been combined in the Affect Infusion Model (Forgas, 1995).

Moods can vary throughout the course of a day in response to small everyday occurrences (L. A. Clark & Watson, 1988; Isen, 1984) and vary from day to day within individuals (L. A. Clark & Watson, 1988). Although moods are transitory, they can lead to fluctuations in self-esteem and influence behaviour (Forgas, 2013). Moreover, the general propensity to experience more PA/NA in everyday life impacts on our more general sense of well-being and is considered to be a stable personality trait (Diener, Suh, Lucas, & Smith, 1999; Watson, Clark, & Tellegen, 1988). Most people in the general population experience more PA than NA (Crawford & Henry, 2004) and have a positive self-concept (Sinclair, et al., 2010). Perhaps as a result, people are generally positively biased in their evaluations of self-related stimuli, such as the letters in their own name (Koole, Dijksterhuis, & van Knippenberg, 2001). Epley and Whitchurch (2008) investigated whether such positivity biases generalise to self-image. They morphed photographs of the participant's face with images of attractive and unattractive composite faces, resulting in a set of faces varying in

attractiveness. Participants, (particularly those with high self-esteem) were more likely, and faster, to select an attractively enhanced version of their face as their own out of line-ups containing their original and morphed image. Verosky and Todorov (2010) found similar effects when the trustworthiness, rather than attractiveness was altered.

Previously, Mirams et al. (2014) investigated whether appearance-based enhancement effects also generalise to judgements of healthiness and whether dispositional tendencies to experience positive versus negative emotion moderate face health judgement. Other studies have found that unhappy participants lack enhancement effects, or even perceive themselves to look worse than they do in reality. Individuals with symptoms of body dysmorphia, for example, have shown reduced perceptions of self-attractiveness (Clerkin & Teachman, 2008). Mirams et al. (2014) developed a novel face health judgement task intended to measure 'healthiness of self-image' that is, whether people perceive themselves to look more or less healthy than they do in reality. Photographs of the participant's face were altered to look more/less healthy, by adding/subtracting the amount of redness in the skin. Red skin coloration is associated with cardiovascular fitness and increasing redness in small amounts increases perceived healthiness and attractiveness of human faces (see Stephen, Coetzee, Smith, & Perrett, 2009). In addition to redness, the amount of yellowness and lightness in the skin (Stephen, Smith, Stirrat, & Perrett, 2009) impacts on judgements of health, as well as facial configuration (e.g., symmetry, Jones, et al., 2001) and facial adiposity (Coetzee, Perrett, & Stephen, 2009). However, we chose to manipulate a single attribute of facial appearance due to our psychophysical method. During the face health judgement task, participants were asked to decide whether they currently felt more or less healthy than each face version (i.e., on each trial, participants decided 'how do I feel compared to this version of my own face?'). Psychophysical methods were used to measure perceptual thresholds (i.e., to identify the face version that approximated participants' self-image). Compared to

participants who self-reported high levels of PA, participants who self-reported high levels of NA, judged themselves as equivalent to greener, unhealthy looking face versions, suggesting that the saw themselves as looking less healthy (Mirams, et al., 2014). This could lead unhappy participants to experience further NA, forming a vicious cycle (Mirams et al., 2014). Indeed, Zell and Balcetis (2012) note that higher level cognition and action are often based on lower-level, perceptual processes.

Although changing skin tone impacts on how attractive, as well as how healthy faces look (Stephen, Coetzee, et al., 2009), in our current study, we were interested primarily in how mood affects judgements of healthiness, given the importance of subjective perceptions of health for individuals' more general sense of well-being (e.g., Diener, et al., 1999) and because evaluations of current health status have implications for health behaviours, including health care utilization and treatment adherence (see Croyle & Uretsky, 1987; Salovey & Birnbaum, 1989). The face health judgement task also has methodological advantages, compared to paradigms used previously, for example Epley and Whitchurch's paradigm. Altering skin tone, rather than morphing the participant's face with attractive or unattractive composite faces, means that faces differ only in how healthy they look, but not in the degree to which they look similar/dissimilar to the participant. Furthermore, the face health judgement task has ecological validity, as people often use skin tone as a cue to health (e.g., Stephen, Coetzee, et al., 2009).

Although previous findings are in line with the idea that mood might alter how people see themselves, there is a lack of direct experimental evidence. It is possible that variables other than mood (e.g., self-esteem, or other personality traits) accounted for individual differences in self-image in Mirams et al.'s (2014) study. Therefore, the first aim of the current study was to more stringently test the hypothesis that mood biases self-image, by investigating the effect of laboratory-induced, rather than pre-existing mood on face health

judgement. As the same participants completed the face health judgement task under three mood conditions, we could be more confident that any differences in face health judgement were due to mood, rather than any individual difference variables. Although previous studies have found that transient moods impact on global, self-reported judgements of health-status (e.g., Croyle & Uretsky, 1987; Howren & Suls, 2011; Salovey & Birnbaum, 1989), the impact of mood on appearance-based judgements of healthiness has never been investigated.

Our second aim was to investigate how emotions affect social comparisons of face health. In Mirams et al.'s (2014) study, participants also completed a version of the face health judgement task in which they judged how they felt compared to healthy and unhealthy looking versions of a stranger, who was matched in age, gender and initial facial redness (i.e., 'how do I feel compared to this version of a stranger's face?'). Participants who self-reported high levels of NA judged themselves as equivalent to less healthy looking versions of a stranger's face and high self-esteem was associated with more favourable social comparisons of health (Mirams et al., 2014).

This finding is in line with previous evidence suggesting that mood might affect social comparisons (for a review see Wheeler & Miyake, 1992). For example, depressed people frequently make upward social comparisons (compare themselves to people who are better off, e.g., Bäzner, Brömer, Hammelstein, & Meyer, 2006; Butzer & Kuiper, 2006). In turn, there is evidence that social comparisons affect our mood; experimentally manipulating upward social comparisons (comparing one's self to an attractive model) increases negative affect (Tiggemann & McGill, 2004), whereas downward social comparisons (comparing ones' self to a person who is worse off), increase positive affect (Gibbons & Gerrard, 1989).

Moreover, social comparisons are thought to shape our self-evaluations (e.g., Festinger, 1954; Tao, Zhang, Li, & Geng, 2012) and even shape our self-image (e.g., Zell & Balcetis, 2012). Using an adapted version of Epley and Whitchurch's paradigm, Zell and

Balcetis (2012) found that after viewing same-gender attractive models, students rated themselves as less attractive and selected a less attractive version of their face as their own out of a line-up. It is possible, therefore, that mood congruent social comparisons (judging oneself favourably in comparison to a stranger when in a positive mood/judging oneself unfavourably when in a negative mood) could exacerbate the effect of mood on self-image.

In the current study, we recruited participants who reported low to moderate levels of PA and NA (who did not have a dispositional tendency to experience high levels of PA or NA), so that we might more easily manipulate their mood, and measured their face health judgement after a positive, negative and neutral mood induction. We predicted that mood would bias healthiness of self-image. After a positive mood induction, we expected that participants would judge themselves as equivalent to healthier looking versions of their own face (compared to after a neutral or negative mood induction), but after a negative mood induction, we expected that the same participants would judge themselves as equivalent to less healthy looking versions of their own face (compared to after a neutral or positive mood induction). After a neutral mood induction, we expected that participants would be relatively unbiased in their face health judgement. We also predicted that mood would affect social comparisons of health, that is, judgements of how healthy one feels in comparison to a stranger's face, in a similar manner. Given that social comparisons shape our self-evaluations (Festinger, 1954; Tao, et al., 2012), we also expected that performing mood congruent social comparisons during the 'stranger' version of the face health judgement task would have a knock on effect to the 'self' version of the task. As a result, there would be a stronger effect of mood on self face health judgement when participants previously performed mood congruent social comparisons during the stranger version of the task.

2. METHOD

2.1 Participants and recruitment

An advertisement for participants with normal or corrected to normal vision, without colour blindness, was placed on the University of Manchester research volunteering website. The advertisement included a link to an online version of the Positive and Negative Affect Schedule (PANAS; Watson, et al., 1988) which consists of ten positive and ten negative feelings and emotions (e.g., active, determined, excited, afraid, distressed, and irritable). Respondents rate the extent to which they have felt each feeling/emotion during the past few weeks on a scale from one (very slightly or not at all) to five (extremely). The PANAS has good construct validity and test-retest reliability and this version provides an indication of dispositional affectivity (Watson, et al., 1988). Out of 86 responses, individuals who scored in the middle (18 – 29) or lower (<17) quartile on the NA scale and who also scored in the middle (29 – 36) or lower (<28) quartile on the PA scale (N = 46) were invited to take part in the main phase of the study. The final sample consisted of 34 participants (29 female; aged 18-27; 91.18% Caucasian, 2.94% South Asian, 5.88% were Black). The effect of facial reddening on judgements of health does not differ according to the ethnicity of the observer or the face being judged (Stephen, Coetzee, et al., 2009). Participants gave informed consent prior to completing the online PANAS and the main experiment, which were approved by the University of Manchester Ethics Committee.

2.2 Overview of study design and procedure

After meeting the experimenter to get their photograph taken, participants attended three, forty five minute testing sessions, spaced one week apart. Figure one illustrates the study design and testing session procedure. In each session, participants completed a state version of the PANAS (indicating to what extent they felt each emotion 'right now') and questionnaires measuring physical symptoms and self-esteem; the Patient Health Questionnaire-15 (PHQ-15; Kroenke, Spitzer, & Williams, 2002) and the Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1989). We previously found scores on these questionnaires

to be associated with performance on the face health judgement task (Mirams et al., 2014). Participants then underwent a positive, negative or neutral mood induction (mood order was counterbalanced between participants, so that a third of participants did the positive mood induction first, a third did the negative mood induction first and a third did the neutral mood induction first). Participants were then informed that they were moving on to the second half of the study and completed the PANAS again, followed by 'self' and 'stranger' versions of the face health judgement task. Self/stranger task order was counterbalanced between participants.

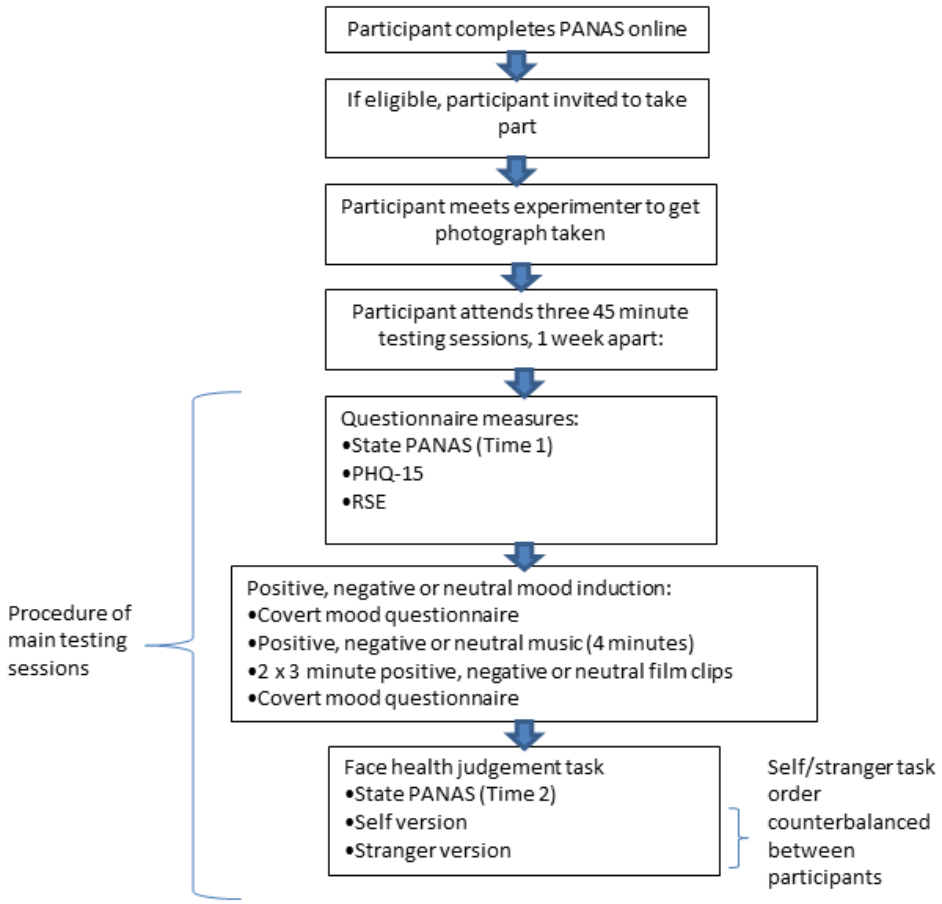

Figure 1. Illustration of the study procedure

2.3 Mood induction materials, procedure and manipulation check

Participants were told that the first part of the study was investigating the effect of listening to music on the perception of film clips. Initially, participants completed a music and film questionnaire which included seven filler items about music and films (e.g., 'how often do you watch films') mixed with five mood items. Participants were asked to rate their overall mood from -10 (very negative) to 10 (very positive) and the extent to which they were currently feeling sad, happy, angry and warm hearted from 1 (not at all) to 7 (very intensely). These items were taken from a modified version of the Differential Emotion Scale (Izard, Dougherty, Bloxom, & Kotsch, 1974) used in previous studies validating the use of emotional film clips to induce mood states (McHugo, Smith, & Lanzetta, 1982; Schaefer, Nils, Sanchez, & Philippot, 2010). Participants then listened to a four minute track of positive, negative or neutral music. Bach's 'Brandenberg Concerto No.3' played by jazz flutist Hubert Laws and Provokiev's 'Russia under the Mongolian yoke' played at half speed were used to induce positive and negative emotions, respectively (Green, Sedikides, Saltzberg, Wood, & Forzano, 2003; Wood, Saltzberg, & Goldsamt, 1990). Following Wood et al. (1990) a Chopin Waltz: 'No. 12 in F minor' was used in the neutral condition.

Next, participants watched two three minute positive, negative or neutral film clips. Following Schaefer et al. (2010) we used clips from Forest Gump (in which a father is reunited with his son) and Dead Poets Society (in which students in a classroom express solidarity with their teacher) to elicit positive affect. To elicit negative affect, we used clips from Schindler's list (which shows dead bodies being carried to mass graves in a concentration camp) and Dangerous Minds (in which a teacher tells her class that a classmate has been killed). These clips elicit positive and negative affect, respectively and increase self-reported arousal (Schaefer, et al., 2010). In the neutral mood condition, we used clips from Blue (in which a man clears out his desk) and The Lover (in which a girl travels in a car), which elicit low self-reported arousal, low PA and low NA (Schaefer, et al., 2010).

Participants were asked to play close attention to the film clips because they would be asked to answer questions regarding their content afterwards.

After viewing the film clips, participants were asked to complete a second music and film questionnaire which included twelve items about the music and the film clips (e.g., 'what were the prisoners wearing?') plus the same five mood items included in the first music and film questionnaire. Participants' ratings of happiness and warm heartedness (and sadness and anger) were summed to produce index scores of positive (and negative) mood as a manipulation check.

2.4 The Face Health Judgement task materials and colour manipulation checks

Before attending the first testing session, participants met the experimenter to get their photograph taken. Following Stephen et al. (2009), the amount of redness/greenness in their skin was then manipulated by adding/subtracting values of A to/from their original photograph in the CIE LAB colour space, using Matlab. The CIE LAB colour space consists of three axes representing the amount of lightness-darkness (L), redness-greeness (A) and bluenenss-yellowness (B) in an image. Details of the photography and image manipulation procedures are described in Mirams et al. (2014). 125 face versions[1] were produced, which varied in how healthy they looked (from -17 to +14 values of A; see Figure two for examples). In a pilot study, reported in Mirams et al., (2014), fifty nine participants (students at The University of Manchester) rated fourteen versions of their own and a stranger's face (ranging from -15 to +10 values of A) on a scale from zero (very unhealthy) to nine (very healthy). Red-tinged face versions were rated as more healthy looking than green-tinged versions, with the highest healthiness ratings for face versions with + 1.5 (*M* rating = 6.19) to +2.5 (*M* rating = 6.41) values of A. On average, the red versions of participants own, and stranger's faces were rated as more healthy looking than the green versions ($t(57)= 5.98$, $p < .001$, $d= 1.58$ and $t(57)= 3.17$, $p =.002$, $d = .84$, respectively).

To determine whether or not our colour manipulation *primarily* affected how healthy faces looked (over and above any other characteristic, such as attractiveness), a separate sample of fifty four naïve participants (students at Liverpool John Moores University) were presented with fourteen versions of a stranger's face (a 19 year old female), ranging from –15 to +10 values of A. Participants viewed all of the face versions on the computer screen, at the same time, and were asked to complete the following sentence, "These face versions differ in how …….. they look". 85.19% of participants chose a health-related word ('healthy', 'unhealthy', 'ill', or 'healthy glow'). 3.7% chose 'attractive', the remaining 11.11% suggested either 'pale', 'photogenic', or 'green'.

If participants gave their consent, their photograph was used to make a stranger face health judgement task to be viewed by other participants. Photographs of nine participants from the current study and fourteen participants from our previous study were presented to participants in the stranger versions of the task. The stranger's face allocated to each participant was decided by matching participants in age, gender and as closely as possible in initial facial redness (A values; *M* difference = -.06) to another face. Out of the twenty three faces allocated to be used in the stranger versions of the task, nine were judged by more than one participant. Post-hoc ratings of attractiveness (from one: not at all attractive, to ten: extremely attractive) by twelve independent raters suggested that participant and stranger faces were equivalent in attractiveness. Attractiveness ratings for the participants' faces ranged from 2.35 to 7.12 (*M* = 4.65, *SD* = 1.30). Ratings for the faces used as strangers ranged from 2.41 to 7.47 (*M* = 4.72, *SD* = 1.34).

2.5 The Face Health Judgement task procedure

Participants were seated approximately 60cm in front of a computer monitor and e-prime software was used to present participants with one of the 125 face versions at a time, starting with a green-tinged, unhealthy looking face (-15 values of A). On each trial, participants were asked, "Do you currently feel more or less healthy than this face?" and were instructed to base their decision on the skin tone. Participants responded by pressing "M" (more) or "L" (less) on the computer keyboard. 50% thresholds were determined using a computerised forced choice adaptive procedure; that is, the face version presented on the each trial depended on the participant's responses on previous trials. Parameter Estimation by Sequential Testing (PEST; Taylor & Creelman, 1967) was used to select the face version on each trial. PEST is an adaptive method of quickly and efficiently estimating psychophysical parameters (for a detailed description of the PEST algorithm see Mirams et al., 2014). For example, if participants responded "more healthy" to the green-tinged face, successively redder faces were presented until the algorithm found the participant's 50% threshold; i.e., the face version to which they responded "more healthy" and "less healthy" equally often (which should roughly correspond to their internal representation of self-image). 50% thresholds > 0 indicate that the participant judges themselves as equivalent to a face with redness added (healthier looking) whereas 50% thresholds < 0 indicate that the participant judges themselves as equivalent to a face version with redness subtracted (green-tinged/less healthy looking). To maintain variability and stop the task becoming too difficult, adaptive staircase trials (75%) were mixed with dummy trials (25%) on which a random face version was selected. The adaptive staircase task took five minutes to complete on average (*M* no. of trials = 77, SD = 50) and participants completed a self and a stranger version. During the self version, participants judged how they felt compared to the different versions of their own face ('how do I feel compared to this version of my own face?'). In the stranger version, participants judged how they felt compared to the different versions of a stranger's face

('how do I feel compared to this version of a stranger's face?'). Whereas the self face health judgement task was intended to measure self-image, the stranger face health judgement task was intended to measure social comparisons, i.e., do people see themselves as equivalent to healthy or unhealthy looking versions of a stranger's face?

3. RESULTS

3.1 Questionnaire measures and mood manipulation check

PHQ-15 and RSE scores in each testing session were not normally distributed and remained so after transformation attempts. Unlike in our previous study (Mirams et al., 2014), scores on each measure were not correlated with 50% thresholds (Spearman's $r \leq .15$, $ps \geq .39$), perhaps due to the narrower range of scores in this sample of participants (*M* range PHQ-15 = 10.00, *M* range RSE = 19.00) compared to the sample in our previous study (*M* range PHQ-15, = 22.00, *M* Range RSE = 26.00).

To investigate whether the mood inductions had the intended effects, overall mood ratings and positive and negative mood index scores were compared before and after the mood induction procedures in each testing session in a series of time (T1: before mood induction/T2: after mood induction) x mood condition (positive/negative/neutral) repeated measures ANOVAs. Significant interactions were followed up by paired samples t-tests (with the significance level lowered to $p = .01$ to control for multiple tests) to determine whether self-reported mood changed significantly from before to after the mood induction in each testing session and whether self-reported mood after the mood induction was significantly different in each testing session/mood condition.

3.2 Overall mood ratings

The initial 2 (time) x 3 (mood condition) ANOVA showed a significant interaction between time x mood ($F(2,66) = 39.77$, $p < .001$). Paired samples t-tests showed that overall mood ratings increased significantly from before ($M = 4.68$) to after ($M = 5.97$) the positive

mood induction ($t(33) = 4.40$, $p < .001$, $d = 1.55$) and from before ($M = 4.21$) to after ($M = 5.21$) the neutral mood induction ($t(33) = 2.87$, $p < .01$, $d = .75$), but decreased significantly from before ($M = 5.23$) to after ($M = .97$) the negative mood induction ($t(33) = -5.84$, $p < .001$, $d = 1.53$). Further paired samples t-tests showed that overall mood ratings at T2 were significantly higher after the positive compared to after the negative mood induction ($t(33) = 6.78$, $p < .001$, $d = 1.78$), tended to be higher after the positive compared to the neutral mood induction ($t(33) = 2.46$, $p = .02$, $d = .65$) and were significantly lower after the negative compared to after the neutral mood induction ($t(33) = -5.97$, $p < .001$, $d = 1.57$).

3.3 Positive mood index

The initial 2 (time) x 3 (mood condition) ANOVA showed a significant interaction between time and mood ($F(2,66) = 36.07$, $p < .001$). Paired samples t-tests showed that positive mood increased significantly from before ($M = 9.62$) to after ($M = 10.32$) the positive mood induction ($t(33) = 3.13$, $p < .01$, $d = .82$), but decreased significantly from before ($M = 9.71$) to after ($M = 6.50$) the negative mood induction ($t(33) = -7.97$, $p < .001$, $d = 2.09$) and tended to decrease from before ($M = 9.21$) to after ($M = 8.68$) the neutral mood induction ($t(33) = -1.81$, $p < .08$, $d = 0.48$). Further paired samples t-tests showed that positive mood at T2 was significantly higher after the positive compared to after the negative mood induction ($t(33) = 7.31$, $p < .001$, $d = 1.92$) and after the positive compared to the neutral mood induction ($t(33) = 5.13$, $p < .001$, $d = 1.35$) and was significantly lower after the negative compared to after the neutral mood induction ($t(33) = -4.91$, $p < .001$, $d = 1.30$).

3.4 Negative mood index

The initial 2 (time) x 3 (mood condition) ANOVA showed a significant interaction between time and mood ($F(2,64) = 13.21$, $p < .001$). Paired samples t-tests showed that negative mood tended to decrease from before ($M = 3.29$) to after ($M = 2.85$) the positive mood induction ($t(33) = -2.27$, $p = .03$, $d = .60$), increased significantly from before ($M =$

3.26) to after (M = 6.82) the negative mood induction (t(33) = 8.07, p < .001, d = 2.12) and did not change significantly from before (M = 3.29) to after (M = 2.97) the neutral mood induction (t(33) = -1.48, p = .15, d = 0.39). Further paired samples t-tests showed that negative mood at T2 was significantly higher after the negative compared to after the positive mood induction (t(33) = 8.79, p < .001, d = 2.31) and after the negative compared to after the neutral mood induction (t(33) = 8.74, p < .001, d = 2.30) but was not significantly different after the positive compared to after the neutral mood induction (t(33) = -.64, p = .52, d = 0.17).

In summary, after the positive mood induction, ratings of overall and positive mood increased, whereas self-reported negative mood decreased. After the negative mood induction, ratings of overall and positive mood decreased, whereas self-reported negative mood increased.

3.5 Did mood affect face health judgement?

A mixed design ANOVA was conducted, with mood (positive, negative and neutral) and task version (self, stranger) as within subjects factors, task order (self or stranger version first) as a between subjects factor and 50% thresholds as the dependent variable. Initial facial redness (M initial redness = 12.79, SD = 1.55) was also included as a covariate. There was a mood x task order interaction (F(2,62) = 4.14, p = .02). There was also a mood x version x task order interaction (F(2,62) = 8.81, p < .001). No other effects were significant (ps $\geq$ .11). To follow up this interaction, two separate repeated measures ANOVAs were conducted for each task order group (self version first, stranger version first).

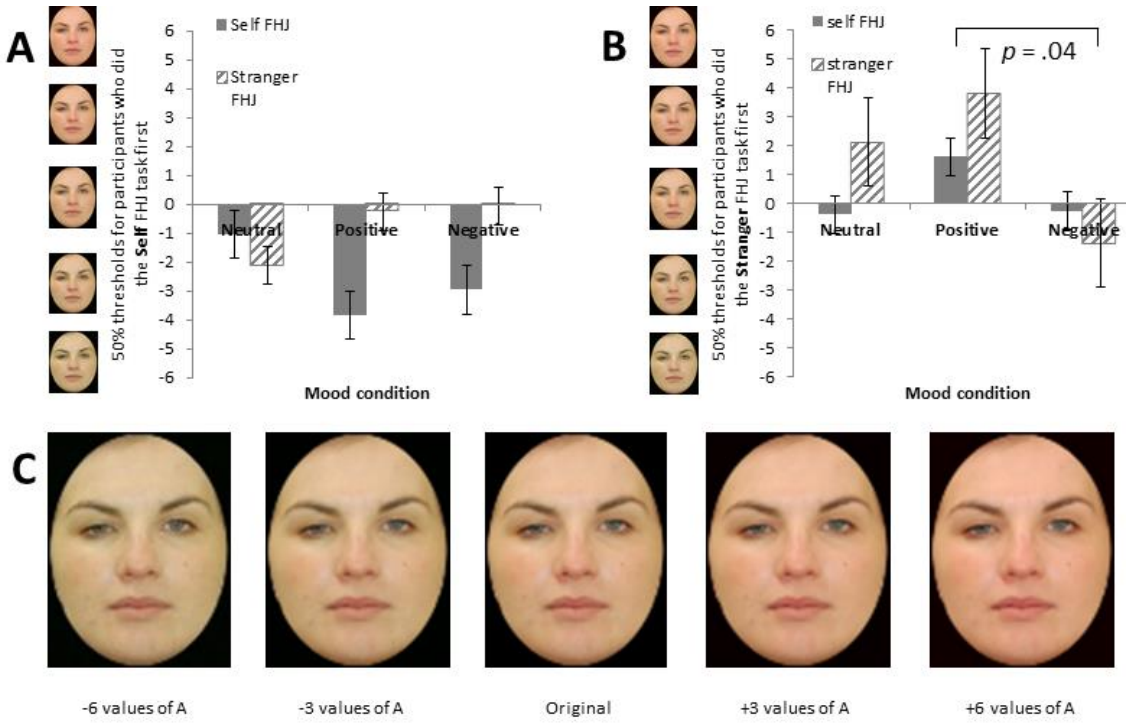

Figure 2. 50% thresholds on the self and stranger versions of the face health judgement task, in each mood condition, for participants who did the self (**A**) or the stranger (**B**) version of the face health judgement task first. The y-axis shows the degree of colour change, in A values, from the original photograph (0). **C** illustrates the colour manipulations for face versions varying from -6 to +6 values of A. *Self 50% thresholds indicate how participants perceived themselves. Stranger 50% thresholds indicate how participants judged themselves in comparison to the stranger. Error bars reflect + 1 standard error of the mean.*

3.5.1 Self version first

For participants who completed the self face health judgement task before the stranger face health judgement task (N = 17), there was no effect of mood (F(2,30) = .49, p = .62) and no interaction between mood and task version (F(2,30) = 2.06, p = .15, see Figure 2A). No other effects were significant (p's $\geq$ .22).

3.5.2 Stranger version first

For participants who completed the stranger version of the face health judgement task before the self version (N = 17), there was a main effect of mood (F(2,30) = 5.77, p = .01, see

Figure 2B). Pairwise comparisons with Bonferroni correction indicated that 50% thresholds were significantly higher in the positive (M = 2.78) than in the negative mood conditions (M = -.53, p = .04), but did not differ between the neutral (M = .84) and positive (p = .46), or neutral and negative conditions (p = .52). That is, when these participants were in a happy compared to unhappy mood their point of indifference fell at a redder, healthier looking face, for both versions of the face health judgement task. There was also a mood x initial facial redness interaction (F(2,30) = 6.89, p = .003), suggesting that the effect of mood differed depending on the amount of redness in participants' original photographs. To follow up this interaction, difference scores (average 50% thresholds after the negative mood induction – average 50% thresholds after the positive mood induction) were correlated with initial facial redness. Low initial facial redness was associated with higher difference scores (r = .67, p = .01), suggesting that mood had more of an effect on face health judgement for participants with less redness in their original photograph. There was no interaction between mood and task version (F(2,30) = .32, p = .73) and no other significant effects (p's $\geq$ .49).

3.6 The association between face health judgement on the self and stranger versions of the task in each task order group

To investigate whether self and stranger face health judgement was more strongly associated for participants who completed the stranger face health judgement task first (as a result of carry-over effects of performing mood congruent social comparisons), a series of correlations were conducted between 50% thresholds on the self and stranger versions of the task in each mood condition, with the significance level lowered to p = .02 to control for thresholds were not associated after the positive mood induction (r = .21, p = .42), but tended to be associated after the neutral (r = .52, p = .03) and negative mood inductions (r = .54, p = .03). For participants who did the stranger version first, self and stranger 50% points were

associated after the positive, negative and neutral mood inductions (r = .58, p = .01, r = .87, p < .001, and r = .85, p < .001, respectively).

4. DISCUSSION

We investigated whether transient mood affects healthiness of self-image and social comparisons of health. Following our previous findings of individual differences in face health judgement in participants with dispositional tendencies to experience positive and negative emotion (Mirams et al., 2014), we expected that when participants were in a negative compared to a positive mood, they would judge themselves as equivalent to less healthy looking versions of their own face and would compare themselves unfavourably against another person. We found mixed support for these hypotheses; when participants initially completed the self face health judgement task (judged themselves against their own face first), mood did not affect face health judgement. However, when participants initially completed the stranger face health judgement task (judged themselves against a stranger's face first), mood did affect face health judgement. This finding is consistent with our prediction that self-stranger task order would moderate the effect of mood on face health judgement, i.e., mood congruent social comparisons would exacerbate the effect of the mood inductions.

Participants who completed the stranger face health judgement task immediately after the mood induction, followed by the self version, judged themselves as equivalent to redder, healthier looking versions of a stranger's face and their own face (with an average of 2.72 values of redness added) after the positive mood induction, compared to after the negative mood induction. When in a negative mood, these participants judged themselves as equivalent to greener, less healthy looking versions of a stranger's face and their own face (with an average of .81 values of redness subtracted). This difference was greater for participants with lower initial facial redness, perhaps because the colour manipulations were

more salient for these participants. This finding corresponds to our previous finding of derogation effects in participants who self-report dispositional tendencies to experience negative moods (Mirams et al., 2014) and is also in line with previous evidence that individuals characterised by NA compare themselves unfavourably to others (e.g., Wheeler & Miyake, 1992). It seems that transient negative mood can also affect self-image and social comparisons of health.

Although other studies have found that negative moods affect self-reported judgements of health status (e.g., Croyle & Uretsky, Salovey & Birnbaum, 1989; Howren & Suls, 2011), our findings suggest that mood also affects appearance-based judgements of health. Unexpectedly, we found no difference in face health judgement when participants were in a neutral compared to a happy, or negative mood, although the differences in 50% thresholds on the stranger version of the task, were in the expected direction (see Figure 2B). Positive biases in self-related judgements when one is in a happy mood, could serve to preserve self-esteem and positive affect (e.g., Crocker & Park, 2003). Altered perceptions of health when one is in a negative mood, on the other hand, could lead people to experience further negative affect and form a vicious cycle (c.f. Mirams et al., 2014). Indeed, other studies have found that people who self-report high (compared to low) levels of NA report more physical symptoms and illnesses without necessarily differing in their objectively measured health (Mora, DiBonaventura, Idler, Leventhal, & Leventhal, 2008; Suls & Howren, 2012; Watson & Pennebaker, 1989; Williams, et al., 2002).

From an evolutionary perspective, it has been argued that persistent (i.e., trait) biases in human judgement and decision making, including biases in social and self-perception, may be adaptive (that is, serve to maximise fitness and the likelihood of survival). For example, Haselton and Nettle (2006) suggest that the fundamental attribution error (making a dispositional attribution for a person's negative behaviour) may have evolved to protect an

individual from the potentially dangerous consequences of failing to detect a harmful, manipulative disposition. Given that there appears to be a reciprocal relationship between positive illusions and physical health, Krebs and Denton (1997) argue that positive illusions may have evolved to help people stay mentally and physically fit (c.f. Snyder & Higgins, 1988). The ultimate functions of mood state-congruent biases in health judgement, are less clear. Although judging one's health favourably when in a happy mood, may serve to maintain that happy mood and in turn a state of good health, judging one's health unfavourably when in a negative mood appears to be maladaptive. Perhaps when we are experiencing negative affect, we momentarily judge ourselves as less healthy, because the occurrence of negative affect may signal a problem/threat to our health. This fits with Affect as Information (Clore, et al., 1994; Schwartz & Clore, 1988) and Affect Infusion Models (Forgas, 1995), according to which we use current feeling states as information to inform our judgements.

Unexpectedly, participants who completed the self face health judgement task immediately after the mood induction, did not show the expected enhancement/derogation effects after the positive/negative mood inductions, on either version of the face health judgement task. Moreover, in contrast to previous findings of enhancement effects in healthy people (e.g., Epley & Whitchurch, 2008; Verosky & Todorov, 2011), these participants did not show enhanced healthiness of self-image in the neutral, or positive mood conditions. This may have been because we pre-selected participants who scored in the low to middle range on the PANAS (i.e., who were less happy to begin with), or because not all participants have an enhanced perception of their self-image. In our previous study (Mirams et al., 2014), participants with high scores on the PA scale were also un-biased in their self face health judgement and in Clerkin & Teachman's (2008) study, students without symptoms of body dysmorphia were accurate in their perceptions of self-attractiveness.

It is also possible that factors other than mood affect face health judgement during the self version of the task, which masked the effects of the mood inductions. Viewing our own, compared to another person's face activates different brain regions, for example (e.g., Platek, et al., 2006) and our own face is a highly emotional stimulus. People often dislike photographs of themselves, perhaps because they do not live up to people's idealised mental representations of themselves (c.f. Epley & Whitchurch, 2006). It is possible that emotional reactions to the photographs, which depended on whether participants liked or disliked the photograph of themselves, may have over-ridden the effects of the mood inductions.

Although mood did not affect performance on the self version of the face health judgement task immediately after the mood inductions, it did affect performance on the self face health judgement task version after participants had completed the stranger version, perhaps because social comparisons exacerbated the effect of the mood inductions. Previous evidence suggests that emotions affect how we judge ourselves against others (e.g., Butzer & Kuiper, 2006; Detweiler-Bedell, Detweiler-Bedell, & Salovey, 2006; Wheeler & Miyake, 1992) and in turn, that such social comparisons affect how we feel (e.g., Festinger, 1954; Gibbons & Gerrard, 1989; Tiggemann & McGill, 2004) and even affect our self-image (Zell & Balcetis, 2012). It is possible that making mood congruent social comparisons during the stranger face health judgement task (judging oneself favourably in comparison to a stranger after the positive mood induction/judging oneself unfavourably against the stranger after the negative mood induction) enhanced positive/negative affect which in turn, impacted on self-image. In line with this idea, performance on the self and stranger versions of the face health judgement task was more strongly associated in participants who completed the stranger version of the task first. That is, performing less favourable social comparisons during the stranger version of the task, was associated with less favourable face health judgement when participants subsequently performed the self version of the task.

The effect of social comparison on self evaluation is likely to be affected by the attractiveness of a stranger's face, i.e., whether they are more/less attractive than one's self (e.g., Zell & Balcetis, 2012). We did not match participants to strangers based on attractiveness, therefore, there is a possibility that small discrepancies in self-stranger attractiveness influenced the results (i.e., comparing oneself to a more/less attractive stranger could affect one's mood). However, we do not think that this could account for our significant effects, given that the participants compared themselves to the same stranger in each of the mood conditions. In future studies, self and stranger faces should be matched in attractiveness to rule out his confound.

We are confident our participants were making judgements based on healthiness in the current study. Previous evidence suggests that our colour manipulation affects perceived healthiness (Mirams et al., 2014) and that participants increase redness in the skin to optimise healthiness in photographs of faces (Stephen, Coetzee et al., 2009). In addition, a separate sample of naïve participants spontaneously suggested that our colour manipulation changed how healthy the faces looked. It is likely that our colour manipulation also affected how attractive our face versions looked (previous findings suggest that skin tone impacts on how attractive, as well as how healthy faces look; Stephen, Coetzee, et al., 2009). Therefore, it will be necessary to investigate the effect of mood on attractiveness judgements, before we can conclude that our results are specific to health-related judgements.

5. CONCLUSIONS

We have shown, for the first time, that transient mood affects appearance based social comparisons of health and in turn, self-image. This finding is consistent with previous evidence suggesting that mood and social comparisons can influence self-image (e.g., Mirams et al., 2014; Clerkin & Teachman, 2008, Zell & Balcetis, 2012). Over a longer period of time, mood related changes in social comparison processes and self-image could

exacerbate positive/negative affect and lead to a positive/negative self-concept. One benefit of our psychophysical methodology, compared to self-report techniques, is that it allows us to measure subtle mood related changes in self-evaluation, even if participants are unaware of any change. Therefore, the face health judgement task could be a useful alternative to self-report techniques as a way of measuring state well-being. Because a stranger's face is a more neutral stimulus compared to our own face, the stranger face health judgement task might provide a cleaner measure of how people feel. As we pre-selected participants with low to medium levels of PA and NA, in order to maximise the effect of the mood inductions, further research is needed before we can generalise our findings to the wider population.

6. FOOTNOTES

1 We chose this asymmetrical range due to previous findings from a pilot study (N = 59, reported in Mirams et al., 2014) that although increasing greenness continues to make faces look less healthy the greener the face becomes, increasing redness above ~5 values of A starts to decrease healthiness, therefore we did not deem it necessary to produce face versions above +14 values if A, as it would be extremely unlikely that a participant would consider a red-tinged face to look healthy beyond this point. We do not think that there is a possibility that the asymmetry in our stimulus variation could have influenced the results, as no participant made judgements about the full range of faces and the adaptive staircase procedure was consistent across mood conditions.

7. ACKNOWLEDGEMENTS

The authors would like to thank Magdalena Kroll for her ideas in the early stages of the project and Unilever Research and Development for funding this research. We also wish to thank Jean-Marc Desirer at Unilever R&D Trumbull (USA) for his helpful comments during the design process and Sarah Davies for the use of her photograph in the figures.

8. REFERENCES

Bäzner, E., Brömer, P., Hammelstein, P., & Meyer, T. D. (2006). Current and former depression and their relationship to the effects of social comparison processes. Results of an internet based study. *Journal of Affective Disorders, 93,* 97-103.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36,* 129-148.

Brown, J. D., & Mankowski, T. A. (1993). Self-esteem, mood, and self-evaluation: Changes in mood and the way you see you. *Journal of Personality and Social Psychology, 64,* 421-430.

Butzer, B., & Kuiper, N. A. (2006). Relationships between the frequency of social comparisons and self-concept clarity, intolerance of uncertainty, anxiety, and depression. *Personality and Individual Differences, 41,* 167-176.

Clark, L. A., & Watson, D. (1988). Mood and the mundane: Relations between daily life events and self-reported mood. *Journal of Personality and Social Psychology, 54,* 296-308.

Clark, M. S., & Isen, A. M. (1982). Toward understanding the relationship between feeling states and social behaviour. In A. Hastorf & A. M. Isen (Eds.), *Cognitive and Social Psychology* (pp. 73-108). New York: Elsevier.

Clerkin, E. M., & Teachman, B. A. (2008). Perceptual and cognitive biases in individuals with body dysmorphic disorder symptoms. *Cognition & Emotion, 22,* 1327-1339.

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of Social Cognition* (Vol. 1, pp. 323-417). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Coetzee, V., Perrett, D. I., & Stephen, I. D. (2009). Facial adiposity: A cue to health? *Perception, 38,* 1700.

Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43,* 245-265.

Croyle, R. T., & Uretsky, M. B. (1987). Effects of mood on self-appraisal of health status. *Health Psychology, 6,* 239-253.

Detweiler-Bedell, B., Detweiler-Bedell, J. B., & Salovey, P. (2006). Mood-congruent perceptions of success depend on self-other framing. *Cognition & Emotion, 20,* 196-216.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125,* 276-302.

Epley, N., & Whitchurch. (2008). Mirror, mirror on the wall: Enhancement in self-recognition. *Journal of Personality and Social Psychology Bulletin, 34,* 1159-1170.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117-140.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117,* 39-66.

Forgas, J. P. (2013). Belief and affect: On the mental pre-cursors of health-related cognition and behaviour. *Journal of Health Psychology, 18,* 3-9.

Gibbons, F. X., & Gerrard, M. (1989). Effects of upward and downward social comparison on mood states. *Journal of Social and Clinical Psychology, 8,* 14-31.

Green, J. D., Sedikides, C., Saltzberg, J. A., Wood, J. V., & Forzano, L. B. (2003). Happy mood decreases self-focused attention. *British Journal of Social Psychology, 42,* 147-157.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology Review, 10,* 47-66.

Howren, M. B., & Suls, J. (2011). The symptom perception hypothesis revised: Depression and anxiety play different roles in concurrent and retrospective physical symptom reporting. *Journal of Personality and Social Psychology, 100,* 182-195.

Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of Social Cognition* (pp. 179-236). Hillsdale, NJ: Lawrence Erlbaum.

Izard, C. E., Dougherty, F. E., Bloxom, B. M., & Kotsch, N. E. (1974). *The Differential Emotions Scale: A method of measuring the meaning of subjective experience of discrete emotions*. Nashville, TN: : Vanderbilt University, Department of Psychology.

Jones, B. C., Little, A. C., Penton-Voak, I. S., Tiddeman, B., Burt, D., & Perrett, D. (2001). Facial symmetry and judgements of apparent health: Support for a "good genes" explanation of the attractiveness–symmetry relationship. *Evolution and Human Behavior, 22*, 417-429.

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80*, 669-685.

Krebs, D. L., & Denton, K. (1997). Social illusions and self-deception: The evolution of biases in person perception. In J. A. Simpson & D. T. Kenrick (Eds.), *Evolutionary social psychology* (pp. 21-48). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine, 64*, 258-266.

McHugo, G. J., Smith, C. A., & Lanzetta, J. T. (1982). The structure of self-reports of emotional responses to film segments. *Motivation and Emotion, 6*, 365-385.

Mirams, L., Poliakoff, E., Zandstra, E. H., Hoeksma, M., Thomas, A., & El-Deredy, W. (2014). Feeling bad and looking worse: Negative affect is associated with reduced perceptions of face-healthiness. *Plos One, 9*, e107912.

Mora, P. A., DiBonaventura, M. D., Idler, E., Leventhal, E. A., & Leventhal, H. (2008). Psychological factors influencing self-assessments of health: Toward an understanding of the mechanisms underlying how people rate their own health. *Annals of Behavioral Medicine, 36*, 292-303.

Platek, S. M., Loughead, J. W., Gur, R. C., Busch, S., Ruparel, K., Phend, N., Panyavin, I. S., & Langleben, D. D. (2006). Neural substrates for functionally discriminating self-face from personally familiar faces. *Human Brain Mapping, 27*, 91-98.

Rosenberg, M. (1989). *Society and the Adolescent Self-Image. Revised edition.* Middletown, CT: Wesleyan University Press.

Salovey, P., & Birnbaum, D. (1989). Influence of mood on health-relevant cognitions. *Journal of Personality and Social Psychology, 57*, 539-551.

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion, 24*, 1153-1172.

Schwartz, N., & Clore, G. L. (1988). How do I feel about it?: The informative function of affective states. In K. Fielder & J. P. Forgas (Eds.), *Affect, Cognition and Social Behavior: New Evidence and Integrative Attempts* (pp. 513–523). Toronto: Hofgrefe International.

Sedikides, C. (1992). Changes in the valence of the self as a function of mood. In M. S. Clark & M. S. Clark (Eds.), *Emotion and social behavior.* (pp. 271-311). Thousand Oaks, CA, US: Sage Publications, Inc.

Simon, H. (1982). Affect and Cognition: Comments. In M. S. Clark & T. Fiske (Eds.), *Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition* (pp. 333-342). Hillsdale, NJ: Lawrence Erlbaum.

Sinclair, S. J., Blais, M. A., Gansler, D. A., Sandberg, E., Bistis, K., & LoCicero, A. (2010). Psychometric properties of the Rosenberg Self-Esteem Scale: Overall and across demographic groups living within the United States. *Evaluation and the Health Professions, 33*, 56-80.

Snyder, C. R., & Higgins, R. L. (1988). Excuses: their effective role in the negotiation of reality. *Psychological Bulletin, 104*, 23.

Stephen, I. D., Coetzee, V., Smith, M. L., & Perrett, D. I. (2009). Skin Blood Perfusion and Oxygenation Colour Affect Perceived Human Health. *Plos One, 4*.

Stephen, I. D., Smith, M. J. L., Stirrat, M. R., & Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *International Journal of Primatology, 30*, 845-857.

Suls, J., & Howren, M. B. (2012). Understanding the physical-symptom experience: The distinctive contributions of anxiety and depression. *Current Directions in Psychological Science, 21,* 129-134.

Tao, R., Zhang, S., Li, Q., & Geng, H. (2012). Modulation of self-esteem in self- and other-evaluations primed by subliminal and supraliminal faces. *Plos One, 7.*

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41,* 782 - 787.

Tiggemann, M., & McGill, B. (2004). The role of social comparison in the effect of magazine advertisements on women's mood and body dissatisfaction. *Journal of Social and Clinical Psychology, 23,* 23-44.

Verosky, S. C., & Todorov, A. (2010). Differential neural responses to faces physically similar to the self as a function of their valence. *NeuroImage, 49,* 1690-1698.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity *Psychological Review, 96,* 234-254.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98,* 219-235.

Wheeler, L., & Miyake, K. (1992). Social comparison in everyday life. *Journal of Personality and Social Psychology, 62,* 760-773.

Williams, P. G., Colder, C. R., Lane, J. D., McCaskill, C. C., Feinglos, M. N., & Surwit, R. S. (2002). Examination of the neuroticism-symptom reporting relationship in individuals with type-2 diabetes. *Personality and Social Psychology Bulletin, 28,* 1015-1025.

Winkielman, P., Knutson, B., Paulus, M., & Trujillo, J. L. (2007). Affective influence on judgments and decisions: Moving towards core mechanisms. *Review of General Psychology, 11,* 179-192.

Wood, J. V., Saltzberg, J. A., & Goldsamt, L. A. (1990). Does affect induce self-focused attention? *Journal of Personality and Social Psychology, 58,* 899-908.

Zell, E., & Balcetis, E. (2012). The influence of social comparison on visual representation of one's face. *Plos One, 7,* e36742-e36742.

FIGURE CAPTIONS

Figure 1. Illustration of the study procedure

Figure 2. 50% thresholds on the self and stranger versions of the face health judgement task, in each mood condition, for participants who did the self (**A**) or the stranger (**B**) version of the face health judgement task first. The y-axis shows the degree of colour change, in A values, from the original photograph (0). **C** illustrates the colour manipulations for face versions varying from -6 to +6 values of A. *Self 50% thresholds indicate how participants perceived themselves. Stranger 50% thresholds indicate how participants judged themselves in comparison to the stranger. Error bars reflect + 1 standard error of the mean.*