Journal of Experimental Social Psychology 71 (2017) 83–95

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

Emotion expression modulates perception of animacy from faces

Natalie C. Bowling *, Michael J. Banissy

Department of Psychology, Goldsmiths College, University of London, UK

HIGHLIGHTS

- Participants rated whether faces appeared animate at points along a morph continuum.
- Animacy perception was influenced by both stimulus and individual variability.
- Male faces were judged to appear more animate than female faces.
- Happy faces were perceived as more animate than neutral faces.
- An externally-oriented thinking style was associated with lower animacy thresholds.

ARTICLE INFO

Article history:
Received 13 July 2016
Revised 13 February 2017
Accepted 14 February 2017
Available online 14 March 2017

Keywords:
Animacy
Face perception
Emotion
Alexithymia
Objectification

ABSTRACT

Discriminating real human faces from artificial can be achieved quickly and accurately by face-processing networks, but less is known about what stimulus qualities or interindividual differences in the perceiver might influence whether a face is perceived as being alive. In the present studies, morphed stimuli differing in levels of animacy were created. Participants made judgements about whether the face appeared animate at different levels along the morph continuum. The faces varied in terms of emotional expression (happy vs. neutral) and gender. Male faces were judged to be animate at a lower threshold (i.e., closer to the inanimate end of the continuum) than female faces. Animacy was also perceived more readily in faces with happy expressions than neutral. These effects were observed across two separate studies involving different participants and different sets of stimuli (animate faces morphed with dolls or those morphed with computer generated faces). Finally, the influence of interindividual variability in personality traits on animacy perception was examined. This revealed that an externally oriented cognitive style, a component of alexithymia, was associated with lower thresholds for perceiving animacy, for animate faces morphed with dolls. The findings are discussed in relation to inter- and intra-individual variability in animacy perception and social interaction.

1. Introduction

The accurate identification of animate (i.e., living beings capable of independent actions, thoughts, and emotions) human faces from inanimate objects is vital for social interaction and carries a key evolutionary advantage. This process relies on perceptual cues from the whole face, including structural and featural information (Balas & Horski, 2012; Balas & Tonsager, 2014). Two features of particular importance are the eyes and mouth (Looser & Wheatley, 2010), which are thought to play a key role in detecting animacy as they convey socially relevant information such as speech, intentions, and emotional expressions (Emery, 2000; Langton, Watt, & Bruce, 2000).

Previous experiments on the perception of animacy have used stimuli that are morphed between human and dolls' faces, and report a threshold for perceiving life at 67% (Looser & Wheatley, 2010). Several studies also compare the 'Point of Subjective Equality' (PSE). This refers to the point on the morph continuum at which stimuli are judged to be 50% animate. A stimulus at this point on the continuum is therefore equally likely to be perceived as either animate or inanimate. This point consistently falls closer to the animate end of the stimulus continuum, and ranges between 56–68% (Balas, 2013; Balas & Horski, 2012; Hackel, Looser, & Van Bavel, 2014; Looser & Wheatley, 2010). Two stimulus factors that have been shown to influence the PSE are the social identity of the stimulus (e.g. Hackel et al., 2014; Swiderska, Krumhuber, & Kappas, 2012) and stimulus gender (e.g. Balas, 2013). With regards to gender, Balas demonstrated that female faces are less likely to be perceived as animate than male faces; and animate faces less likely to be perceived as female than male. These results have led to suggestions that they reflect the dehumanisation or objectification

---

* Corresponding author at: Department of Psychology, Goldsmiths College, University of London, New Cross, London SE14 6NW, UK.
E-mail address: n.bowling@gold.ac.uk (N.C. Bowling).

http://dx.doi.org/10.1016/j.jesp.2017.02.004

of the faces of women (Balas, 2013), but this has not yet been empirically tested for animacy perception. A purely perceptual account of why gender of a face may influence animacy judgments can also be made. Female faces are associated with narrower jawlines and lighter skin pigmentation compared with male faces (Brown & Perrett, 1993; Frost, 1988) and as such share a closer similarity with the doll faces typically used in animacy experiments than male faces do. This may result in female face stimuli being rated as less animate than male stimuli. These differing hypotheses are not mutually exclusive, and it is possible that both processes contribute to the reported gender differences in animacy perception. Here we sought to investigate both the objectification and perceptual hypotheses of gender difference. Objectification is addressed by incorporating a measure of the extent to which participants objectify women's bodies. If objectification underlies the gender differences seen in animacy perception, then participants who demonstrate greater objectification of women should show higher animacy thresholds for female faces than those who score low on objectification of women.

In addition to gender and social identity a number of other factors may be important to animacy perception. For example, face perception research frequently uses achromatic stimuli to avoid confounding effects of differences in facial pigmentation. However, the majority of previous facial animacy perception studies (Balas, 2013; Hackel et al., 2014; Looser & Wheatley, 2010; Powers, Worsham, Freeman, Wheatley, & Heatherton, 2014; Swiderska et al., 2012; Wheatley, Weinberg, Looser, Moran, & Hajcak, 2011) have used chromatic stimuli. It is therefore unclear how important colour cues such as skin pigmentation are for detecting animacy in a face. This casts some doubt on the validity of comparing animacy thresholds for different stimulus faces (e.g. gender and racial groups) where colour cues have not been controlled. Where achromatic stimuli have been used (Balas & Koldewyn, 2013; Balas & Tonsager, 2014; Looser, Guntupalli, & Wheatley, 2013), there has been no direct comparison of achromatic and chromatic stimuli, and so it remains unclear whether the results can be applied to animacy judgements made with chromatic stimuli, either in previous studies or real-world perception.

Further, to our knowledge, no previous research on animacy perception has considered the effect of the emotional expression of the stimulus on animacy thresholds. The studies mentioned above have averaged together ratings for several different stimuli, regardless of the emotion expressed. Given the social significance of emotion expression (Keltner & Kring, 1998), it seems likely that this factor may influence animacy perception from faces. More specifically, if animacy reflects a capacity for experiencing emotion (Looser et al., 2013), it follows that a face expressing emotion would be more likely to be perceived as animate than a face with neutral expression. The importance of the eyes and mouth in making animacy judgements (Looser & Wheatley, 2010) lends further support to this hypothesis, since these features are also particularly relevant for conveying and perceiving emotion (Eisenbarth & Alpers, 2011; Emery, 2000; Langton et al., 2000; Yuki, Maddux, & Masuda, 2007). Collectively, this evidence indicates that emotional expression may influence animacy perception judgements.

In conjunction with properties of the stimulus, individual differences in the observer can influence animacy perception judgements. For example, the readiness with which facial animacy is perceived has recently been linked to the desire for social connection. In this study by Powers et al. (2014), scores on a Need to Belong Scale (NTBS; Leary, Kelly, Cottrell, & Schreindorfer, 2013) correlated with animacy perception thresholds, such that individuals with a greater desire for social acceptance and belonging perceived animacy at a lower threshold. Further, participants subjected to an experimental manipulation to induce feelings of social disconnection also judged animacy to occur at a lower threshold than those who received a 'socially connected' induction. The authors proposed that these results reflect an adaptive strategy on the part of individuals who feel socially isolated, where perceiving animacy more readily increases the likelihood of valuable social interaction. This idea ties in with the suggestion that animacy is perceived more readily for in-group members than out-group due to a greater motivation for social interaction with the in-group (Hackel et al., 2014). If attributing animacy to an ambiguous stimulus indeed reflects a strategy to gain social interaction, then thresholds should also be lower for individuals with increased loneliness. Epley, Akalis, Waytz, and Cacioppo (2008) report that self-reported loneliness correlates positively with mental state attribution in objects. In this study more lonely individuals were more likely to describe an inanimate object as having "a mind," "intentions," and "emotions." Further, experimentally induced social disconnection led to greater attribution of anthropomorphic traits related to social connection to their pets. As yet the relation between loneliness and animacy perception has not been explored.

Previous results linking desire for social connection with increased animacy perception (Powers et al., 2014) suggest that other social factors may also influence how animacy is perceived. Individuals with high trait social anxiety appear to demonstrate attentional biases towards socially relevant stimuli. However, the direction of this bias is not yet clear, with increased attention observed in certain contexts, and avoidance in others (for reviews see Bögels & Mansell, 2004; Heinrichs & Hofmann, 2001). In either case we might predict socially anxious individuals to demonstrate altered detection of animacy in human faces, compared with controls. In the case of increased attention, individuals with social anxiety may identify animacy more readily, leading to lower animacy thresholds; and in the case of avoidance individuals may be less likely to detect animacy in the face, leading to higher thresholds. Evidence from Epley et al. (2008) favours the former hypothesis, demonstrating that experimentally induced fear leads to greater likelihood of perceiving faces in ambiguous line drawings, compared with induced social disconnection. The hypothesised relation between social anxiety and animacy perception therefore provides an interesting research question, as well as a tool for understanding the cognitive biases associated with the condition.

A final trait factor that could be implicated in the detection of animacy is alexithymia. Alexithymia is a subclinical personality trait reflecting difficulties identifying and describing emotions, and the tendency to focus attention externally, while reducing emotional experiences (Bagby, Parker, & Taylor, 1994). It is reported in higher levels in males than females (Franz et al., 2008). Alexithymia is thought to involve deficits in processing emotion information (Lane et al., 1996) and in facial emotion recognition specifically, though studies of the latter have so far yielded mixed results (Cook, Brewer, Shah, & Bird, 2013; Grynberg et al., 2012; Jongen et al., 2014; Pandey & Mandal, 1997; Parker, Taylor, & Bagby, 1993). If, as we hypothesise, emotion recognition is involved in detecting animacy, then it may follow that individuals high in alexithymia show differential facial animacy processing, compared with those who score low. Alexithymia has also been associated with impairments in empathy (Bird et al., 2010; Parker, Taylor, & Bagby, 2001) and in 'mentalizing,' understanding the mental states of others (Moriguchi et al., 2006). Since animacy perception involves making a judgement about whether a stimulus has the capacity to possess mental states, this provides further support for the notion that alexithymia would be associated with reduced perceptions of animacy.

1.1. Current study

With the aforementioned studies in mind, the current online study compared the effect of stimulus qualities and individual differences of the perceiver on perception of animacy in ambiguous face stimuli. These stimuli were created by morphing images of human faces with visually matched doll faces that varied in colour (achromatic vs. chromatic), gender (male vs. female) and emotional expression (happy vs. neutral). The influence of individual differences in the perceiver relevant to social interaction on animacy judgements was also assessed. The relation between interindividual variability in the following traits and facial animacy perception were examined: 'Need to Belong' (as per Powers et al., 2014), loneliness, social anxiety, alexithymia and objectification. Specifically, we made the following predictions:

1. Male faces would be judged to appear animate at a lower threshold than female faces.
2. Faces displaying emotion would be perceived to be animate at a lower threshold than faces with neutral expression.
3. Facial animacy judgements would rely on colour cues in the face, such as skin pigmentation.
4. Perception of animacy would be influenced by individual variability in traits including need to belong, loneliness, social anxiety, alexithymia and objectification.

2. Experiment one

The first experiment aimed to address each of the above hypotheses, using a novel stimulus set formed of doll-human morphs.

2.1. Method

2.1.1. Participants

The target sample size was 90. This target was calculated using an a priori power analysis for a within-subjects *t*-test with 0.8 power and 0.05 alpha level, based on the effect size previously obtained by Hackel et al. (2014; Cohen's $d = 0.3$) when comparing within-subjects animacy perception judgements of two stimulus types. 105 participants were then recruited online using the University College London Sona System, to account for some attrition from the online task. Volunteers were given a £7.50 Amazon voucher for completing the study. Data from one participant was excluded as the individual completed the study twice, leaving 104 participants (49 female, 55 male, age range 18–39 years, $M = 26.6$, $SD = 6.7$). Ethnicity was classified into Asian ($N = 43$), Black ($N = 5$), Caucasian ($N = 48$), or mixed/other ethnic background ($N = 8$). One participant chose not to complete the '*alive*' rating task, resulting in 103 participants for this section only. Some participants also missed or chose to omit items on the self-report scales, meaning that overall scores could not be calculated. This resulted in only 102 completing participants for the Need to Belong and Loneliness Scales, and 103 participants for 'Difficulty Identifying Feelings' and total scores on the Toronto Alexithymia Scale. No further data was collected following analysis for these participants.

2.1.2. Materials and methods

2.1.2.1. Stimuli. Face stimuli were created by morphing together images of human faces from the Radboud Faces Database (RaFD; Langner et al., 2010) with images of dolls, using FantaMorph software (Version 4; Abrosoft Co., Beijing, China). Dolls were selected to represent male and female faces, with happy and neutral expressions (Male, neutral $N = 6$; Male, happy $N = 3$; Female neutral $N = 4$; Female happy $N = 6$). All stimuli (both dolls and human faces) were Caucasian and human faces wore no cosmetics, piercings, facial hair or other distinguishing features. Stimuli were 596 × 736 pixels and displayed in an oval frame, removing external features (hair, ears, neck, etc.; see Fig. 1a). All stimuli are available at http://dx.doi.org/10.5281/zenodo.204416. They can also be obtained by emailing the corresponding author.

Three rating tasks were completed. In each task still images were selected from each morph at 10% intervals, creating 11 still images representing different levels of animacy for each morph, and 209 stimuli in total per block. For animacy threshold judgements, stills were selected at 2% intervals, creating 50 images for each of the 19 morphs.

2.1.2.2. Procedure. The online experiment comprised five main sections, running as follows. All questionnaire measures and experimental tasks are listed here:

1. Self-report questionnaires (see details below for information on each)
   a. Demographic information
   b. Need to Belong Scale
   c. UCLA Loneliness Scale
   d. Toronto Alexithymia Scale
   e. Social Interaction Anxiety and Social Phobia Scales (short versions)
   f. Objectification Questionnaire (Male)
   g. Objectification Questionnaire (Female)
2. Rating Task 1: *Whether the face appears to be alive*
3. Animacy threshold judgements
4. Rating Task 2: *Whether the face is able to feel pain*
5. Rating Task 3: *Whether the face has a 'mind'*

All participants completed the tasks in the order above. The order of rating tasks followed that used by Looser and Wheatley (2010). As the authors suggest, ratings of whether the face had a mind were blocked last, to avoid this influencing other animacy judgements. The threshold task was given after the 'alive' rating task to avoid influencing these ratings, since a similar judgement is being made, but mid-way through the longer rating task blocks to prevent participant fatigue. All tasks followed the procedure used by Looser and Wheatley. On starting the experiment, participants were randomly assigned to either the chromatic ($N = 53$) or achromatic ($N = 51$) condition. In the chromatic condition stimuli were viewed in their original colour, and in the achromatic condition stimuli were seen at 0% saturation.

2.1.2.2.1. Animacy thresholds. To set thresholds for perceived animacy, participants were able to scroll through each morph at 2% intervals. They were asked to "scroll along until you find the point where you think the face changes from *having the appearance of being alive* to *not having the appearance of being alive*. Then select the first image on the *alive* side of that threshold" (see Fig. 1b for an example). Faces were seen as fully inanimate when the scroll bar was on the left, and animate when on the right. Starting positions of the scrollbar were randomised so that they first appeared at each end point on 50% of trials.

2.1.2.2.2. Rating tasks. For each rating task participants were required to make a response on the given criteria for each of the 209 stimulus faces. Responses were given on a 7-point Likert scale, with 1 meaning definitely appears alive/feels pain/has a 'mind' and 7 meaning definitely doesn't appear alive/feel pain/have a 'mind'. Each face was shown on screen for 500 ms, after which participants responded by pressing 1–7 on their keyboard (see Fig. 1c for an example trial).

2.1.2.3. Self-report measures. Individual personality differences were assessed with a battery of self-report questionnaires, these are detailed below.

2.1.2.3.1. Need to Belong Scale. In an attempt to replicate the findings of Powers et al. (2014), the *Need to Belong Scale* (NTBS; Leary et al., 2013) was used as a measure of desire for social connections. Participants were required to rate the extent to which 10 statements, such as "I do not like being alone," were characteristic of them, using a 5-point scale ranging from 'Not At All' to 'Extremely.' Possible scores range from 10–50, with a higher score indicating greater desire for social connection. Leary and colleagues report good construct validity for the scale as well as good reliability ($\alpha = 0.81$). Internal consistency was also good in our sample ($\alpha = 0.83$).

2.1.2.3.2. UCLA Loneliness Scale. To further assess the effect of social connection, the *UCLA Loneliness Scale* (Version 3; Russell, 1996) was used. Participants were shown 20 statements, such as "I lack companionship" and asked to indicate how often they felt the way being described, on a 4-point scale from 'Never' to 'Often.' Possible scores range from 20–80, with a higher score representing greater loneliness.

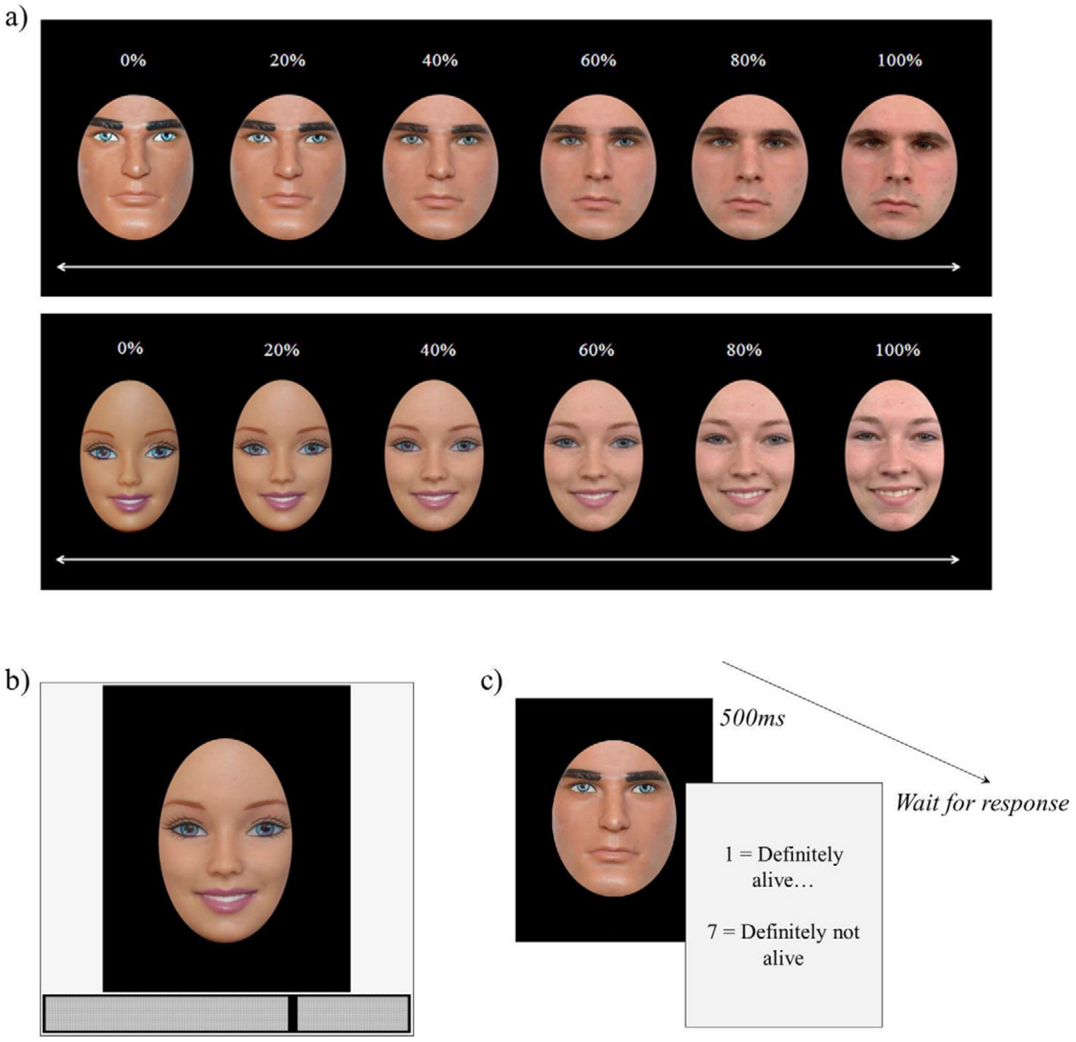

Fig. 1. a) Example male/neutral (upper row) and female/happy (lower row) morph stimuli used in Experiment One. Stimuli are shown here at 20% intervals along the morph continuum. b) Example trial on the threshold task. Participants could use the slider to move up and down the morph continuum at 2% intervals, to select the threshold at which the face first appeared to be animate. c) Example trial on the rating task. Stimuli were displayed for 500 ms before a response was given using the 1–7 number keys.

The scale has been shown to have excellent reliability ($\alpha > 0.90$), as well as construct and convergent reliability. Internal consistency was also excellent in our sample ($\alpha = 0.91$).

2.1.2.3.3. Social Interaction Anxiety and Social Phobia Scales. Short versions of the *Social Interaction Anxiety and Social Phobia Scales* (SIAS-6 & SPS-6; Peters, Sunderland, Andrews, Rapee, & Mattick, 2012) were administered to account for other possible influences on motivation for social interaction. Each scale comprises six statements, and as for the NTBS, participants are required to rate the extent to which each is characteristic of them, using a 5-point scale ranging from "Not At All" to "Extremely". Statements for the SIAS-6 included "I have difficulty making eye contact with others," and for the SPS-6 included "When in an elevator I am tense if people look at me." Each scale generates a score between 0 and 24, with a higher score indicating greater anxiety. Peters and colleagues demonstrate that the validity of these measures is not sacrificed in the shortened versions. In our sample, good internal consistency was found for both SIAS ($\alpha = 0.82$) and SPS ($\alpha = 0.85$).

2.1.2.3.4. Toronto Alexithymia Scale. The *Toronto Alexithymia Scale* (TAS; Bagby et al., 1994) requires participants to indicate the extent which they agree with each of 20 statements, including "I often don't know why I am angry," on a 5-point scale from "Strongly Disagree" to "Strongly Agree." Overall scores can range between 20 and 80, with a higher score indicating more alexithymic traits. Responses can be grouped into three subscales, measuring 'Difficulty Describing Feelings' (5 items), 'Difficulty Identifying Feelings' (7 items), and 'Externally Oriented Thinking' (8 items), which refers to a tendency to focus attention outwards rather than inwardly and includes items such as "I prefer to just let things happen rather than to understand why they turned out that way". The authors report good reliability ($\alpha = 0.81$) as well as validity for the scale. Our sample also reports good internal consistency ($\alpha = 0.83$).

2.1.2.3.5. Self-Objectification Questionnaire. Objectification of men and women was assessed separately with modified versions of the *Self-Objectification Questionnaire* (Noll & Fredrickson, 1998; Strelan & Hargreaves, 2005). In this task participants were required to rank 10 qualities in order of importance, first for men and then for women. These included 5 appearance-based, such as 'physical attractiveness,' and 5 competence-based traits, such as 'physical coordination'. Ranks for competence items can be deducted from appearance items, to obtain an overall objectification score between −25 and 25, with a higher score representing increased objectification. Noll (1996, as cited in Noll & Fredrickson, 1998) reports that the Self-Objectification Questionnaire demonstrates acceptable construct validity.

2.3. Results

2.3.1. Self-report measures

Interindividual self-report questionnaires demonstrated a wide range of responses in all the measured constructs. Means and range of scores are reported in Table 1, and intercorrelations between the scales are displayed in Table 2. Moderate to strong positive correlations were found between subscales of the TAS, between the social anxiety measures SIAS and SPS and between male and female objectification scores. In addition, a significant positive correlation was found between NTBS and male objectification. Those with a greater need to belong reported greater objectification of men. The same relation was not found for female objectification; however, this scale showed a significant positive correlation with the externally oriented thinking (EOT) subscale of the TAS. Individuals who objectified women more also report a more externally oriented cognitive style. Loneliness showed a significant positive correlation with both SIAS and SPS, with more lonely individuals reportedly more socially anxious. Loneliness, SIAS and SPS all correlated positively with total TAS scores, as well as the 'Identifying' and 'Describing' subscales. Increased loneliness and social anxiety appears to be associated with a difficulty labelling identifying and describing emotions. However, of the three, only SPS resulted in a significant positive correlation with EOT.

Prior to examining the influence of trait differences on animacy perception, the influence of participant demographic (age and gender) on the trait measures were examined. Contrary to prior research (Franz et al., 2008), male and female participants showed no significant difference in overall alexithymia scores ($t[101] = 0.30$, $p = 0.762$, Cohen's $d = 0.06$). However, male participants received significantly higher scores on the EOT subscale ($t[102] = 2.49$, $p = 0.014$, Cohen's $d = 0.49$). No gender differences were found on the 'Identifying Feelings' ($t[101] = -1.21$, $p = 0.230$, Cohen's $d = 0.24$) or 'Describing Feelings' subscales ($t[102] = -0.09$, $p = 0.929$, Cohen's $d = 0.02$). Gender differences were found in NTBS ($t[100] = -3.72$, $p < 0.001$, Cohen's $d = 0.74$) and male objectification ($t[102] = -3.01$, $p = 0.003$, Cohen's $d = 0.59$) with female participants demonstrating higher scores than males in both cases. Female objectification did not significantly differ across male and female participants ($t[102] = 1.25$, $p = 0.215$, Cohen's $d = 0.24$). No further gender differences were found on the remaining measures (SIAS: $t[102] = -1.26$, $p = 0.210$, Cohen's $d = 0.25$; SPS: $t[102] = -1.80$, $p = 0.076$, Cohen's $d = 0.35$; Loneliness: $t[100] = -0.26$, $p = 0.794$, Cohen's $d = 0.04$). The effect of participant age on trait measures was significant only for female objectification (see Supplementary Table 1). Objectification scores were negatively correlated with age, such that younger participants scored higher than older participants ($r[102] = -0.23$, $p = 0.018$). Inspection of this effect in each gender group showed that this effect was driven by a highly significant correlation in the female participant group only ($r[47] = -0.39$, $p = 0.006$), with a nonsignificant effect in the male group ($r[53] = -0.09$, $p = 0.521$).

2.3.2. Threshold task

The point on the morph continuum (ranging from 0–100% human) at which the participant reported that the stimulus face first appeared to be animate on the threshold judgement task was combined for each of the 19 stimuli, to calculate a mean animacy threshold for each participant ($M = 68.70$, $SD = 10.96$). To analyse the effect of the gender and emotional expression of the stimulus on perceived animacy thresholds, mean thresholds were also calculated for each gender/emotion group, i.e., male/neutral ($M = 68.29$, $SD = 12.82$), male/happy ($M = 63.63$, $SD = 13.96$), female/neutral ($M = 71.36$, $SD = 12.78$), female/happy ($M = 69.86$, $SD = 11.76$).

Table 1
Means, standard deviations and minimum and maximum scores observed for each of the self-report measures.

Self-report measure	n	M	SD	Range
NTBS	102	29.4	6.7	15–44
Loneliness	102	41.6	10.0	21–69
SIAS	104	12.1	4.5	6–28
SPS	104	10.5	4.4	6–25
Objectification				
Male	104	−5.1	12.1	−25–25
Female	104	5.1	10.9	−25–25
TAS				
Identifying Feelings	103	15.7	6.1	7–30
Describing Feelings	104	13.1	4.1	5–23
Externally Oriented Thinking	104	19.3	4.7	9–30
Overall score	103	48.1	11.5	24–74

2.3.2.1. Interindividual variability in animacy perception.

2.3.2.1. Interindividual variability in animacy perception. To examine how interindividual variability on traits of interest influenced animacy perception, scores on each of the self-report questionnaires were correlated with mean animacy thresholds. Pearson's correlation coefficients are reported in Table 3. All correlations were nonsignificant, with the exception of the EOT subscale of the TAS, which showed a negative correlation with animacy thresholds ($r[101] = -0.29$, $p = 0.003$). A more externally oriented cognitive style was associated with a lower threshold for perceiving animacy, closer to the inanimate end of the continuum (Supplementary Fig. 1). To identify whether this relation was consistent for all subgroups of face stimulus, EOT was correlated with animacy thresholds in each group individually. Significant negative correlations were found for all stimulus subgroups (male/neutral: $r[101] = -0.26$, $p = 0.008$; female/neutral: $r[101] = -0.24$, $p = 0.013$; female/happy: $r[101] = -0.33$, $p = 0.001$) except male/happy stimuli, although this correlation showed a negative trend ($r[101] = -0.13$, $p = 0.204$).

A further analysis was conducted to assess the effect of participant demographics (gender and ethnicity) on animacy threshold judgements. Ethnicity was originally grouped into one of four broad categories: Asian, Black, Caucasian, and mixed/other ethnic background. Since there were relatively few participants with black or mixed/other ethnic background, only Asian and Caucasian groups were compared in this analysis ($N = 91$, 20 Asian males, 23 Asian females, 27 Caucasian males, 21 Caucasian females). A 2 (gender) × 2 (ethnicity) × 2 (stimulus gender) mixed ANOVA was therefore used. Stimulus gender was added as a third factor to identify any interaction effects between the gender of the participant and stimulus. The analysis revealed a main effect of stimulus gender ($F[1,87] = 25.82$, $p < 0.001$, $\eta_p^2 = 0.23$), but no significant main effects or interactions for participant gender or ethnicity (see Supplementary Table 2). The effect of stimulus gender was consistent across male and female participants. Participant gender and ethnicity variables were therefore removed from the remaining analyses.

2.3.2.2. Effects of stimulus variability on animacy perception. The effect of stimulus qualities on threshold judgements was analysed with a 2 (stimulus gender) × 2 (emotion type) × 2 (chromatic condition) mixed ANOVA, with stimulus gender and emotion as within-subjects factors, and chromatic condition as the between-subjects factor ($N = 104$). Main effects of both stimulus gender and emotion type were found, with thresholds for male faces closer to the inanimate end of the continuum than female faces ($F[1,101] = 42.04$, $p < 0.001$, $\eta_p^2 = 0.29$), and thresholds for happy faces closer to the inanimate end than neutral ($F[1,101] = 13.88$, $p < 0.001$, $\eta_p^2 = 0.12$). A further interaction effect was found between stimulus gender and emotion ($F[1,101] = 5.00$, $p = 0.027$, $\eta_p^2 = 0.05$). Post hoc *t*-tests indicate that happy faces were judged to be alive at a significantly lower threshold than neutral faces, for male stimuli only ($t[102] = 3.88$, $p < 0.001$, Cohen's $d = 0.38$), and not for female stimuli ($t[102] = 1.56$, $p = 0.121$, Cohen's $d = 0.16$). A significant gender difference was observed in both neutral ($t[102] = -3.11$, $p = 0.002$, Cohen's $d = 0.30$) and happy stimuli ($t[102] = -6.07$, $p < 0.001$, Cohen's $d = 0.62$) (see Fig. 2). All main effects and interactions involving chromatic condition were nonsignificant (see Supplementary Table 3).

2.3.3. Rating tasks

Participants' ratings for whether the stimulus face '*appears to be alive*,' '*is able to feel pain*' and '*has a mind*' were first subject to a linear

Table 2
Pearson's coefficients for intercorrelations between the self-report measures.

Self-report measure	1	2	3	4	5	6	7	8	9	10
1. NTBS	–									
2. Loneliness	−0.08	–								
df	100									
3. SIAS	0.07	0.51***	–							
df	102	102								
4. SPS	0.05	0.31**	0.62***	–						
df	102	102	104							
Objectification										
5. Male	0.22*	−0.02	0.17	0.13	–					
df	102	102	104	104						
6. Female	0.08	−0.11	0.04	0.02	0.45***	–				
df	102	102	104	104	104					
TAS										
7. Identifying Feelings	0.05	0.40***	0.34***	0.31**	−0.04	−0.02	–			
df	101	101	103	103	103	103				
8. Describing Feelings	−0.01	0.43***	0.47***	0.38***	−0.00	−0.03	0.70***	–		
df	102	102	104	104	104	104	103			
9. Externally Oriented Thinking	−0.12	0.12	0.10	0.21*	−0.09	0.23*	0.23*	0.25*	–	
df	102	102	104	104	104	104	103	104		
10. Overall score	−0.03	0.41***	0.39***	0.39***	−0.05	0.07	0.87***	0.83***	0.62***	–
df	101	101	103	103	103	103	103	103	103	

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

transformation to convert scores from a 1–7 Likert scale to a score between 0 and 1, with 1 representing most animacy, and 0 least animacy. Scores were then combined as for the threshold values, to give a mean rating for each participant, at each animacy level, in each gender/emotion group and overall.

Mean ratings across participants were fitted with a single-term Gaussian function in the curve fitting toolbox for Matlab, using the following equation:

$$f(x) = ae^{\left[-\left(\frac{x-b}{c}\right)^2\right]}$$

As shown in Tables 4–6, a good fit was achieved for all models. This allowed the Point of Subjective Equality (PSE) to be calculated, reflecting the point on the morph continuum at which stimuli were judged to appear 50% animate.

2.3.3.1. Alive ratings. As found on the threshold judgement task, PSEs highlight that male faces were judged to appear alive at lower morph levels (i.e., when 'less human') than female faces, and happy faces at a lower level than neutral faces.

Fig. 3(a) demonstrates that male stimuli were judged to appear more alive than female stimuli at the majority of morph levels, but this difference is not apparent when stimuli are 0% human or 80%–90% human. Holm-Bonferroni-corrected *t*-tests allowed for paired comparisons between ratings for male and female stimuli at each of the 11 animacy levels (results are displayed in Supplementary Table 4). This analysis confirmed that the gender difference in '*alive*' ratings was significant only for morph stimuli between 10 and 70% human, and also just reached significance at 100% human ($t[102] = -2.63$, $p = 0.010$, Cohen's $d = -0.22$), although at this level female faces were judged to appear more animate than male faces.

Fig. 3(b) illustrates a greater effect of emotion occurring at higher morph levels, i.e., when the stimuli were more human. This pattern is reflected in Holm-Bonferroni-corrected *t*-tests, which demonstrate a significant emotion effects at 40% human and above (see Supplementary Table 4).

2.3.3.2. Mind ratings. PSEs for the 'mind' task (shown in Table 5) follow the same pattern as the 'alive' task, demonstrating that participants thought male stimuli and happy stimuli were perceived to 'have a mind' at a lower animacy level than for female or neutral stimuli.

Table 3
Pearson's correlations between self-report measures and animacy threshold judgements.

Self-report measure	Animacy threshold	
	df	*r*
NTBS	99	−0.04
Loneliness	99	0.09
SIAS	101	0.04
SPS	101	0.02
Objectification		
Male	101	−0.12
Female	101	0.04
TAS		
Identifying Feelings	100	−0.03
Describing Feelings	101	0.06
Externally Oriented Thinking	101	−0.29**
Overall score	100	−0.15

** $p < 0.01$.

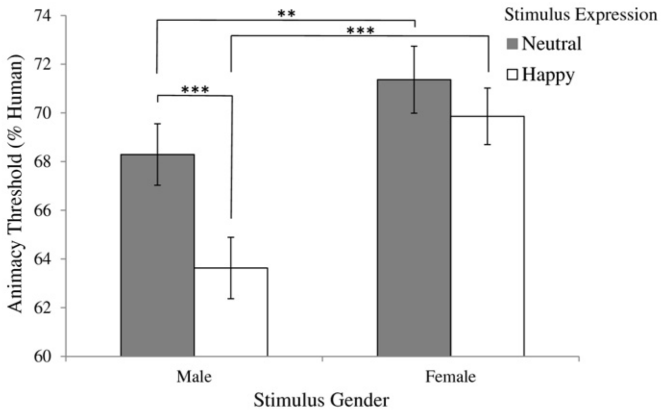

Fig. 2. Mean animacy thresholds for each stimulus subtype, according to emotional expression and gender. Male stimulus faces were perceived to be animate at a lower threshold than female faces, across both emotion groups. Animacy thresholds for male faces were also lower for stimuli with a happy than neutral expressions ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). Error bars represent ±1 S.E.M.

Table 4
PSEs, model parameters and fit statistics (adjusted r-square and root mean square error) for each stimulus group on the 'Alive' rating task.

Stimulus group	PSE (%)	a	b	c	$\overline{R}^2$	RMSE
Male	60.0	0.82	9.88	5.61	0.989	0.031
Female	64.9	0.84	9.95	4.81	0.989	0.033
Happy	58.1	0.87	9.71	5.24	0.990	0.032
Neutral	66.6	0.80	10.21	5.18	0.985	0.036
Male happy	50.6	0.87	9.14	5.49	0.988	0.034
Female happy	61.6	0.87	9.82	4.89	0.990	0.033
Male neutral	63.9	0.81	10.20	5.50	0.984	0.036
Female neutral	70.2	0.78	10.06	4.56	0.986	0.034
Overall	62.4	0.83	9.99	5.25	0.988	0.032

As shown in Fig. 3(d), the gender difference in ratings to whether the face 'has a mind' follow the same pattern as 'alive' ratings, with male faces receiving higher ratings and differences decreasing as the stimulus becomes more human. Overall, the gender difference appears less pronounced than for the 'alive' ratings. Nevertheless, Holm-Bonferroni-corrected *t*-tests between ratings for male and female stimuli on the '*mind*' task report a significant gender difference between 0 and 70% human (shown in Supplementary Table 5).

Similar to the gender difference data, emotion effects follow a similar pattern to ratings on the 'Alive' task, with larger differences observed for more human stimuli (Fig. 3(e)). Again, the differences appear less pronounced on this task than for the 'Alive' task, but Holm-Bonferroni-corrected *t*-tests report significant differences at the same animacy levels as the previous task, at 40% and above (see Supplementary Table 5).

2.3.3.3. Pain ratings. As for previous tasks, PSEs on the pain task indicate that male faces are judged to be animate at a lower level (less human) than female faces. However, in contrast with results from the threshold task and PSE analysis on the 'Alive' and 'Mind' tasks, in this task happy faces were judged 'able to feel pain' at a higher threshold than neutral faces.

Fig. 3(g) highlights that in the case of the pain task, a more consistent gender effect appears, at all morph levels, rather than being more evident at mid-low morph levels. Holm-Bonferroni-corrected *t*-tests comparing the gender difference confirm a significant effect at every level (see Supplementary Table 6).

The effect of emotional expression on stimulus ratings on the pain task also show a different pattern to the alive and mind tasks (Fig. 3(h)). In this case a greater difference is observed at either end of the morph continuum than in the mid-range. This is confirmed by Holm-Bonferroni-corrected *t*-tests, which confirm a significant emotion effect at 0–30% human and 90–100% human only see Supplementary Table 6. The graph also illustrates that neutral faces were judged as being more 'able to feel pain' than happy faces.

2.4. Discussion

In Experiment One, we demonstrate that the perception of animacy in stimuli morphed between human and doll faces is influenced by the gender and emotional expression of the stimulus face, but does not appear to rely on colour cues present in the stimulus. We also show that externally oriented thinking is associated with reduced animacy perception thresholds.

Table 5
PSEs, model parameters and fit statistics (adjusted r-square and root mean square error) for each stimulus group on the 'Mind' rating task.

Stimulus group	PSE (%)	a	b	c	$\overline{R}^2$	RMSE
Male	59.8	0.85	97.63	51.73	0.988	0.034
Female	63.1	0.85	97.14	46.84	0.993	0.026
Happy	58.9	0.88	95.46	48.84	0.993	0.026
Neutral	64.1	0.83	99.15	49.24	0.986	0.036
Male happy	54.8	0.89	92.84	50.14	0.990	0.033
Female happy	61.0	0.87	96.46	47.69	0.995	0.024
Male neutral	62.5	0.84	100.20	52.21	0.983	0.039
Female neutral	66.3	0.82	97.67	44.85	0.990	0.031
Overall	61.5	0.85	97.42	49.18	0.991	0.030

Table 6
PSEs, model parameters and fit statistics (adjusted r-square and root mean square error) for each stimulus group on the 'Pain' rating task.

Stimulus group	PSE (%)	a	b	c	$\overline{R}^2$	RMSE
Male	66.1	0.82	11.97	7.63	0.971	0.041
Female	71.2	0.74	11.06	6.26	0.974	0.039
Happy	69.3	0.68	9.87	5.31	0.986	0.030
Neutral	68.5	1.10	15.51	9.76	0.962	0.047
Male happy	64.6	0.69	9.52	5.43	0.986	0.029
Female happy	71.4	0.68	9.98	5.16	0.985	0.030
Male neutral	66.8	1.09	15.46	9.92	0.961	0.048
Female neutral	70.9	1.09	15.39	9.39	0.961	0.048
Overall	69.0	0.78	11.54	6.96	0.973	0.040

The effect of stimulus gender on perceived animacy in this study supports the findings of Balas (2013). Balas proposed that the effect of stimulus gender may be driven by objectification of female faces. The present study provides the first attempt at testing this hypothesis. Scores on a scale of female objectification showed no significant correlation with animacy perception thresholds, and thus our results do not support this account. In view of this, an alternative perceptual account of gender difference in animacy perception should also be considered. Human female faces share a closer similarity to the doll stimuli used in animacy experiments than male faces, due to narrower jawlines and lighter skin pigmentation than males (Brown & Perrett, 1993; Frost, 1988). This could be the factor that leads to increased animacy thresholds for female stimulus faces.

Male stimuli were also judged to appear more 'alive' and more likely to have a 'mind' than female stimuli at the majority of morph levels, but not when stimuli were 80% human or above. This is unlikely to represent a ceiling effect, since ratings do not reach 100% (consistent with Looser & Wheatley, 2010). The lack of gender difference at the more human end of the morph continuum can possibly be explained by differences in use of cosmetics. All human faces used to create the experimental stimuli wore no make-up. However, many of the female dolls used gave the appearance of wearing make-up, including lipstick and eyeliner, where male dolls did not. These cosmetics can increase the local contrast of the eyes and lips, and in this case could have affected female stimuli at the inanimate end of the spectrum (Balas, 2013). This could lead to female morphed stimuli appearing less realistic than male stimuli at the same animacy level. It is interesting to note that in Balas' (2013) study following the same procedure, human faces were also photographed without removing cosmetics (i.e., cosmetics were present for both doll and human faces), this may explain why a consistent gender difference in animacy judgements was found across the morph continuum.

Experiment Two aimed to test this alternative perceptual explanation of gender differences in animacy perception, by creating a set of stimuli which overcame these potential biases.

3. Experiment Two

In Experiment Two the threshold and rating tasks from the first experiment were repeated using morphed stimuli that combined animate human faces with computer generated inanimate faces. This removed the issue of the exaggeration of feminine facial features and make-up cues present in doll faces, and so allowed an assessment of whether the observed gender differences in Experiment One were influenced by these perceptual factors.

3.1. Method

3.1.1. Participants

Target sample size for the second experiment was 103, to match those collected in Experiment One. 100 participants who did not take

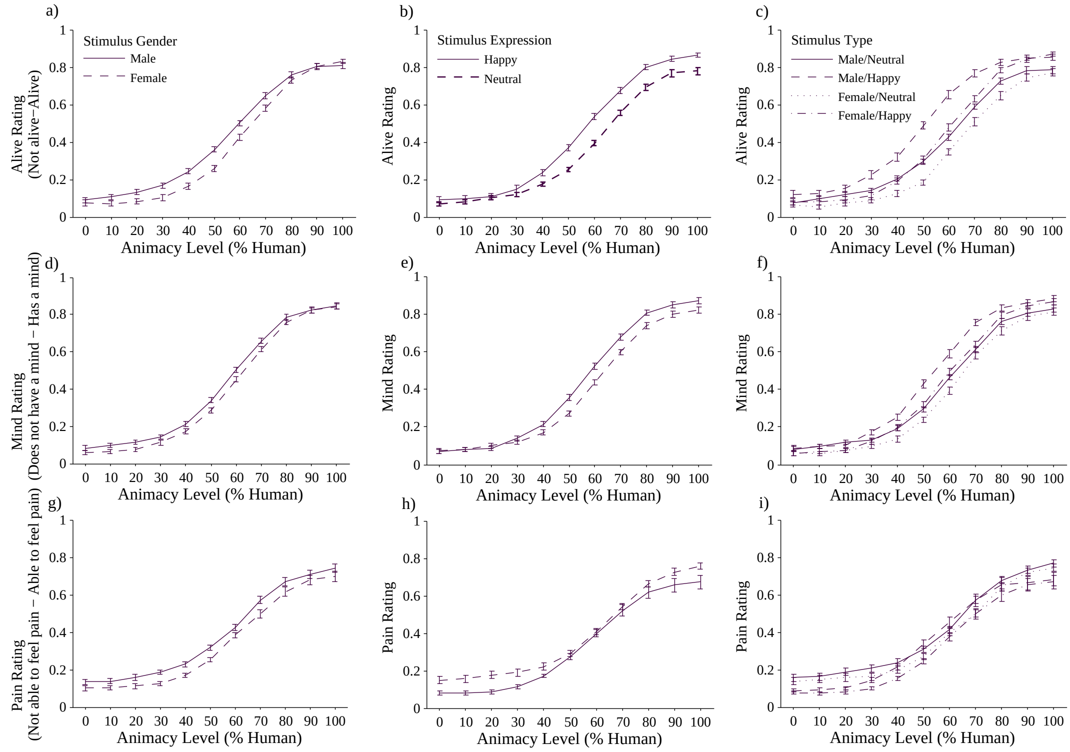

Fig. 3. Mean ratings from Experiment One of (a–c) "whether the face appears to be alive", (d–f) "whether the face has a mind" and (g–i) "whether the face is able to feel pain" at each level of animacy, from 0% human/100% doll to 0% doll/100% human. Y-axis shows ratings from 0 (completely inanimate, e.g. definitely not able to feel pain) to 1 (completely animate, e.g. definitely able to feel pain). Ratings are shown for (a,d,g) each stimulus gender, (b,e,h) each stimulus emotion, and (c,f,i) each stimulus gender/emotion group. In both emotion groups, male stimuli were perceived to be more alive, more likely to have a mind, and more able to feel pain than female stimuli, from 0 to around 80% human. In both gender groups, happy stimuli were judged to appear more alive, and more likely to have a mind than neutral stimuli, from around 40% human. In contrast, neutral faces were perceived to be more able to feel pain than happy faces, at both extremes of the continuum. Error bars represent ±1 S.E.M.

part in Experiment One (30 female, 70 male, age range 18–61 years, $M = 29.2$, $SD = 9.2$) were recruited online from the website Prolific.ac. Ethnicity was classified into Asian ($N = 30$), Black ($N = 4$), Caucasian ($N = 60$), or mixed/other ethnic background ($N = 6$). All participants were paid £5 for taking part.

3.1.2. Materials and methods

3.1.2.1. Stimuli. For Experiment Two, a new set of stimuli were created using FaceGen Modeller (Singular Inversions, Toronto, Canada). FaceGen was used to create computer generated inanimate versions of human faces from the RaFD (Langner et al., 2010). These inanimate versions were then morphed with the originals to make a set of stimuli highly controlled to match features across the morph continuum (see Fig. 4). As in Experiment One, stimuli were selected to represent four categories: male, neutral expression ($N = 4$), male happy expression ($N = 4$), female neutral expression ($N = 4$), female happy expression ($N = 4$). Faces were all Caucasian and without make up or other distinguishing features, and were displayed in a frame removing external features. All stimuli were 400 × 400 pixels. All stimuli are available at

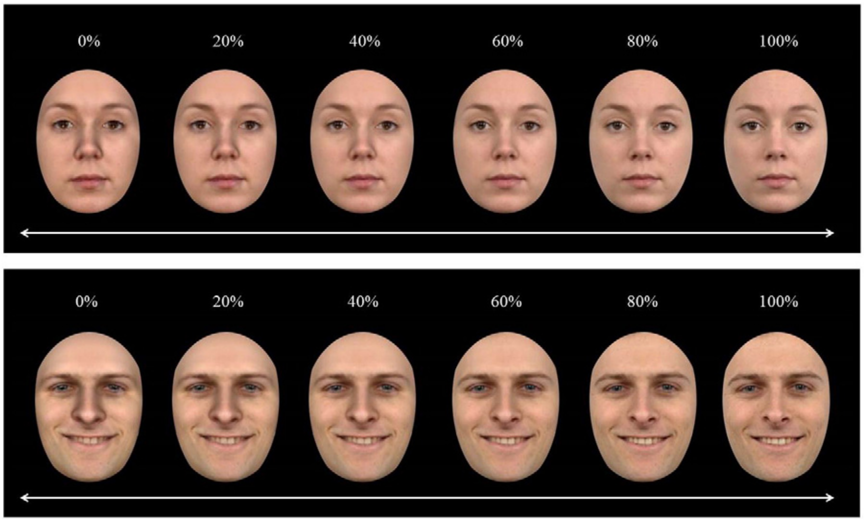

Fig. 4. Example stimuli as used in Experiment Two. The upper row depicts female/neutral stimuli and the lower row male/happy. Stimuli are shown here at 20% intervals along the morph continuum.

http://dx.doi.org/10.5281/zenodo.204453. They can also be obtained by emailing the corresponding author.

3.1.2.2. Procedure. Participants completed all tasks as described in Experiment One, in the same order, with two exceptions. Since in the previous experiment the externally oriented thinking subscale of the TAS (Bagby et al., 1994) was the only trait variable found to be correlated with animacy thresholds, only this questionnaire measure was retained for the second experiment. Additionally, as colour condition (chromatic vs. achromatic) did not significantly influence animacy judgements in the previous experiment, stimuli were displayed in colour for all participants.

3.3. Results

3.3.1. Self-report measures

Internal consistency on the TAS (Bagby et al., 1994) was again shown to be good in this second sample of participants ($\alpha = 0.83$). Overall scores on the TAS ($M = 51.09$, $SD = 11.97$), and scores on the Describing Feelings ($M = 14.29$, $SD = 4.35$), Identifying Feelings ($M = 17.24$, $SD = 6.23$), and externally oriented thinking ($M = 19.56$, $SD = 4.45$) subscales were calculated for each participant. Firstly, individual differences in TAS scores were analysed in terms of age and gender. Pearson's correlation analyses showed no significant relation between age and overall TAS scores ($r[98] = -0.14$, $p = 0.165$), or with externally oriented thinking ($r[98] = 0.07$, $p = 0.498$), but scores on the Describing Feelings subscale did reach significance ($r[98] = -0.20$, $p = 0.043$), and for Identifying Feelings was marginally significant ($r[98] = -0.18$, $p = 0.079$). In both cases TAS scores decreased with age. Gender differences in TAS scores were compared with independent *t*-tests. This revealed no significant gender differences in overall scores ($t[99] = 1.05$, $p = 0.298$, Cohen's $d = 0.22$) or any of the subscales (Identifying Feelings: $t[99] = 0.92$, $p = 0.363$, Cohen's $d = 0.19$; Describing Feelings: $t[99] = 1.45$, $p = 0.151$, Cohen's $d = 0.30$; Externally oriented thinking: $t[99] = 0.12$, $p = 0.904$, Cohen's $d = 0.03$).

3.3.2. Threshold task

Mean threshold judgements were calculated for all stimuli ($M = 58.96$, $SD = 15.76$), and for each stimulus category, including male/neutral ($M = 58.46$, $SD = 17.67$), male/happy ($M = 59.06$, $SD = 21.71$), female/neutral ($M = 59.64$, $SD = 17.02$), and female/happy faces ($M = 58.67$, $SD = 21.57$).

Comparison of threshold judgements with those made in Experiment One showed heterogeneity of variances between the two groups, according to Levene's test ($F[1,201] = 8.78$, $p = 0.003$). With degrees of freedom adjusted for unequal variances, an independent samples *t*-test showed that thresholds for perceiving animacy were significantly lower (closer to the inanimate end of the continuum) in Experiment Two than in Experiment One ($t(176) = 5.10$, $p < 0.001$, Cohen's $d = 0.73$). This indicates that the FaceGen faces used to create the second set of stimuli appeared more animate than the previous doll faces, shifting thresholds closer to the centre of the morph.

3.3.2.1. Interindividual variability in animacy perception. Scores on the TAS (Bagby et al., 1994) were correlated with mean animacy thresholds to identify the relation between these variables. Pearson's correlation coefficients were not significant for total scores ($r[98] = -0.04$, $p = 0.705$), or for the Describing Feelings ($r[98] = -0.01$, $p = 0.978$), Identifying Feelings ($r[98] = -0.10$, $p = 0.309$) or externally oriented thinking subscales ($r[98] = 0.04$, $p = 0.668$). Animacy perception thresholds do not appear to be related to traits of alexithymia in this sample.

As in Experiment One, the effect of participant age, gender and ethnicity on animacy thresholds was analysed. Pearson's correlation coefficient between participant age and mean threshold was not significant ($r[98] = -0.16$, $p = 0.104$), indicating that threshold judgements were consistent across age groups. An ANOVA was carried out to compare thresholds across male and female participants, and across ethnic backgrounds. Firstly, participants were grouped into categories based on ethnicity, resulting in Asian, Black, Caucasian and mixed/other background. As in Experiment One, Asian (male $N = 21$, female $N = 9$) and Caucasian (male $N = 41$, female $N = 19$) participants represented the largest groups, and were compared in this analysis due to the very limited number of participants in other categories ($N = 10$). A 2 (participant ethnicity) × 2 (participant gender) × 2 (stimulus gender) mixed ANOVA showed no significant main effects of ethnicity ($F[1,86] = 0.46$, $p = 0.499$, $\eta_p^2 = 0.01$) or participant gender ($F[1,86] = 0.51$, $p = 0.479$, $\eta_p^2 = 0.01$), or interaction between the two ($F[1,86] = 1.25$, $p = 0.267$, $\eta_p^2 = 0.01$). There were also no significant main effects or interactions with stimulus gender (see Supplementary Table 7). The results suggest that participants perceived similar thresholds for animacy whether they had the same or other gender and ethnicity to the stimulus face.

3.3.2.2. Effects of stimulus variability on animacy perception. To identify the effect of stimulus gender and emotional expression on threshold judgements a 2 × 2 repeated-measures ANOVA was carried out comparing responses for male and female, and happy and neutral stimuli. Neither the effect of stimulus gender ($F[1,99] = 0.15$, $p = 0.702$, $\eta_p^2 < 0.01$) or emotion ($F[1,99] = 0.01$, $p = 0.919$, $\eta_p^2 < 0.01$) were significant in this analysis. The interaction term was also nonsignificant ($F[1,99] = 0.63$, $p = 0.428$, $\eta_p^2 = 0.01$). Threshold judgements appear to be fairly consistent across stimulus categories.

3.3.3. Rating tasks

Following the procedure for Experiment One, ratings of the extent to which the stimulus '*appears to be alive*,' '*is able to feel pain*' and '*has a mind*' were linearly transformed to give a score between 0 and 1, with 1 representing most, and 0 least animacy. Mean ratings in each of the stimulus and emotion categories, at each animacy level were then calculated for each participant. Data from three participants was excluded from the 'mind' task prior to analysis, as they had given either the maximum or minimum rating possible to every stimulus face within one or more gender/emotion categories. This was taken as an indication of a technical difficulty or fatigue on the task. This resulted in 100 participants for the 'alive' and 'pain' rating tasks, and 97 participants on the 'mind' task only (30 female, 67 male, age range 18–61 years, $M = 29.3$, $SD = 9.4$).

As shown in Fig. 5, ratings on all three dimensions showed relatively little change across the animacy morph continuum. This meant that PSE values could not be calculated from the data. Comparison of mean ratings at each animacy level are therefore described below.

3.3.3.1. Alive ratings. Ratings of the extent to which the stimulus face appeared to be alive were compared in a 2 (gender) × 2 (emotion) × 11 (animacy level) repeated-measures ANOVA (see Supplementary Table 8). This revealed significant main effects of gender ($F[1,99] = 31.90$, $p < 0.001$, $\eta_p^2 = 0.24$), with male faces rated as appearing more alive than female faces, and emotion ($F[1,99] = 200.41$, $p < 0.001$, $\eta_p^2 = 0.67$), with happy faces appearing more alive than neutral faces. The main effect of animacy level was also significant ($F[1,99] = 21.51$, $p < 0.001$, $\eta_p^2 = 0.18$), with faces rated as appearing more alive towards the animate end of the continuum, although it should be noted that this effect was smaller than that of emotion or gender. The analysis also revealed a significant interaction between gender and emotion ($F[1,99] = 10.34$, $p = 0.002$, $\eta_p^2 = 0.10$), and between gender and animacy ($F[10,990] = 2.76$, $p = 0.002$, $\eta_p^2 = 0.03$). Holm-Bonferroni-corrected post hoc *t*-tests demonstrate that the effect of gender on ratings for happy faces was significant only at the 70% level of animacy ($t[99] = 3.23$, $p = 0.002$, Cohen's $d = 0.35$). However, the gender effect was significant for neutral faces at 0–50 and 70–80% human (shown in

Supplementary Table 9), indicating that the effect of gender was present more for neutral faces, towards the inanimate end of the continuum.

3.3.3.2. Mind ratings. Ratings for whether the stimulus face appeared to have a mind also showed a main effect of gender ($F[1,96] = 5.04$, $p = 0.027$, $\eta_p^2 = 0.05$), with higher ratings associated with male stimuli than female, and emotion ($F[1,96] = 54.22$, $p < 0.001$, $\eta_p^2 = 0.36$), with higher ratings given to stimuli with happy expressions than neutral. The main effect of animacy level was also significant, with ratings increasing across the morph continuum ($F[10,960] = 32.46$, $p < 0.001$, $\eta_p^2 = 0.25$). In this task no significant interactions between the three variables were found (see Supplementary Table 8).

3.3.3.3. Pain ratings. For the rating task requiring participants to decide whether the stimulus face was capable of experiencing pain, again a main effect of gender ($F[1,99] = 16.69$, $p < 0.001$, $\eta_p^2 = 0.14$), and emotion ($F[1,99] = 51.50$, $p < 0.001$, $\eta_p^2 = 0.34$) were observed. However, in this case the effects were reversed compared with previous tasks, where a greater capacity to experience pain was attributed to neutral faces than happy, and to female faces than to male. The main effect of animacy was also significant ($F[10,990] = 3.11$, $p = 0.001$, $\eta_p^2 = 0.03$). Again, no significant interaction effects were shown (see Supplementary Table 8).

3.4. Discussion

The results of Experiment Two partially replicate the findings from Experiment One and show that the effects of stimulus gender and emotional expression on animacy perception are consistent across varying stimulus types. In this regard, the hypothesis that gender differences in animacy judgements are driven by exaggeration of facial features and apparent cosmetics present in female dolls was not supported. When participants were required to rate the extent to which stimulus faces appeared to be alive, or to have a mind, female faces were again perceived as less animate using highly controlled computer generated stimuli without the bias of feminised facial features or apparent cosmetics. This reaffirms a small but robust effect of gender in animacy perception, although the cause of this effect remains unclear.

One may note that the effects of gender and emotion were only found on the rating tasks, and not on the threshold setting task in which participants could freely select the point at which animacy was first perceived. The realistic nature of the inanimate stimuli used in Experiment Two may suggest a reason for the lack of gender and emotion effects on this task. Thresholds were strikingly consistent across each gender and emotion category, with animacy first perceived when the stimulus was approximately 58–59% human. We show that thresholds for the morphed FaceGen stimuli were significantly lower (closer to the inanimate end of the continuum) than for the morphed doll stimuli in Experiment One. This is perhaps unsurprising, since programs such as FaceGen aim to create a more realistic face stimulus than the dolls used

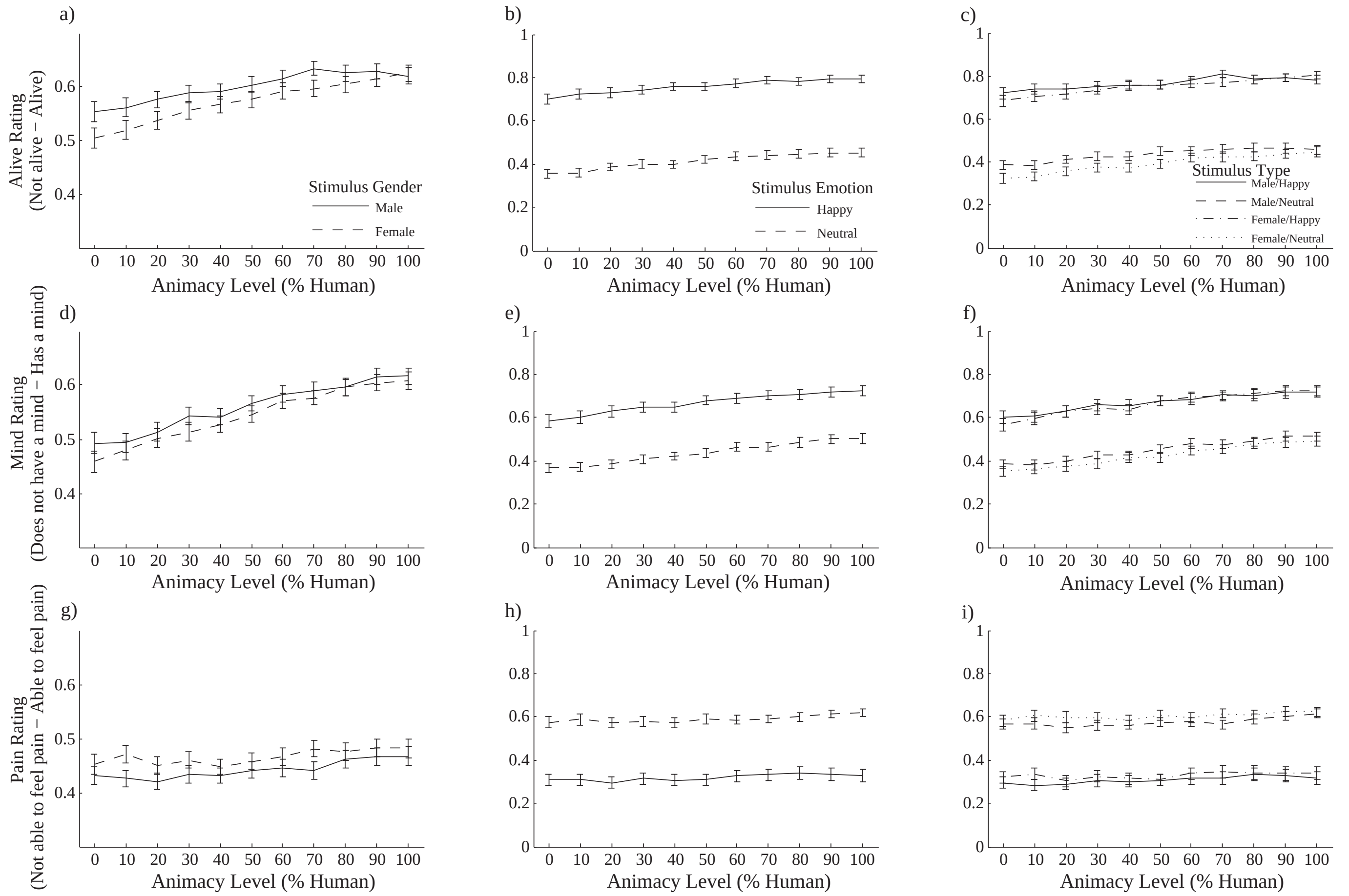

Fig. 5. Mean ratings from Experiment Two of (a–c) "whether the face appears to be alive", (d–f) "whether the face has a mind", and (g–i) "whether the face is able to feel pain", at each level of animacy from 0% human/100% computer generated to 0% computer generated/100% human. Y-axis shows ratings from 0 (completely inanimate, e.g. definitely not able to feel pain) to 1 (completely animate, e.g. definitely able to feel pain). Ratings are shown for (a,d,g) each stimulus gender, (b,e,h) each stimulus emotion, and (c,f,i) each stimulus gender/emotion group. Male faces were perceived as more likely to be alive and to have a mind than female faces, but (in contrast with Experiment One) less likely to be able to feel pain. Faces with happy expression were also perceived as more likely to be alive and to have a mind than neutral faces, but less likely to be able to feel pain. Error bars represent ± 1 S.E.M.

in the previous experiment. A more realistic stimulus at the inanimate end would therefore shift thresholds closer to this end of the continuum. As can be observed from the rating task results (see Fig. 5), stimuli were rated fairly consistently across each level of animacy. This indicates that it may have been difficult for participants to distinguish between the animacy levels, and therefore to select the point at which animacy was first perceived. Participants may then have been more likely to select a similar threshold on every trial, closer to the middle of the continuum, if they found the stimuli too similar to distinguish between.

In the case of rating whether the stimulus face was able to feel pain, effects of gender and emotion were also observed in Experiment Two. However, female stimuli were rated as more able to feel pain (i.e., more animate) than male faces, in contrast with the 'alive' and 'mind' tasks, and with the results of Experiment One. This effect may have arisen from pre-existing assumptions regarding gender differences in sensitivity to pain. Research using the Gender Role Expectations of Pain (GREP) questionnaire (Robinson et al., 2001) demonstrates that both male and female participants rate women as more sensitive to, and less enduring of pain than men. Perhaps this gender bias influenced rating judgements on the task. While this bias did not appear to influence 'pain' judgements in the first experiment, perhaps this discrepancy can also be attributed to differences in the perceived animacy of the two stimulus sets. If participants found it difficult to distinguish faces at different animacy levels in Experiment Two, then perhaps they were more likely to rely on gender information than animacy to make rating judgements.

The relation between higher levels of externally oriented thinking and lowered animacy perception thresholds was not replicated in Experiment Two. This does not seem to be a result of the range of TAS scores obtained from the second participant group, as this was consistent with those in Experiment One. Again, since this correlation measure was based on threshold judgements, the relation between variables may have failed to arise in this case if participants found the stimuli too difficult to distinguish between. Further research is necessary to clarify the link between animacy perception and externally oriented thinking.

4. General discussion

The present research aimed to elucidate the extent to which facial animacy judgements are influenced by a) the gender of the face, b) the emotional expression of the face, c) colour cues and d) interindividual variability relevant to social interaction, across varying stimulus sets. The results demonstrate for the first time that the emotional expression of the stimulus face modulates animacy perception, with happy male faces perceived as being animate at a lower level than neutral male faces. Our results also provide further evidence that animacy is perceived more readily in male faces than in female faces. This effect was not driven by objectification of female faces or by cosmetic features associated with inanimate female stimuli that have been used in the past. Animacy judgements were also not affected by whether the stimulus was chromatic or achromatic, but were found to correlate with an externally oriented cognitive style. Together the results provide important implications for understanding variability in animacy perception and social interaction.

4.1. Stimulus variability factors that contribute to animacy judgments

The mean animacy perception threshold across all stimuli in Experiment One was 69%. This lies very close to the 67% threshold reported by Looser and Wheatley (2010). The sigmoid function observed in animacy ratings across the morph continuum, and PSE values shifted towards the animate end of the continuum also align with results from previous studies (Balas, 2013; Balas & Horski, 2012; Hackel et al., 2014; Looser & Wheatley, 2010). The similarity between animacy thresholds and PSE values in this study and previous studies indicates a reliable and consistent animacy boundary across experimental procedures, and testing environments (either online or in the laboratory). However, animacy thresholds using morphed computer generated stimuli in Experiment Two were significantly lower, around 58% human. This suggests that threshold values previously reported may be limited to stimuli morphed from dolls or statues, as have been consistently used in previous literature. Research should now be extended to other types of inanimate stimuli to identify the extent to which this perceptual threshold may vary.

The hypothesis that animacy perception from faces would be enhanced for stimuli expressing emotion was supported in both experiments. It appears that emotional expressions indicate a capacity to experience emotions. This evidence provides an additional suggestion of why the eyes disproportionately influence animacy judgements, as they convey information about emotional state (Eisenbarth & Alpers, 2011; Yuki et al., 2007). The role of emotional expression highlights the importance of social cues, in addition to the featural and structural properties of the face, in perceiving animacy. This finding does come with a caveat, as the current study used only happy and neutral expressions. It therefore cannot be determined whether the observed effect was a product of emotion expression in general; or of positively-valenced emotion specifically. Future research should aim to establish the role of positive and negative emotion in attributing animacy to faces.

Across both experiments we support previous evidence of the effect of stimulus gender in animacy perception (Balas, 2013), with male faces tending to appear more animate than female faces. We extend these findings to show that both male and female participants judge male faces to appear more animate than female. We found no support for Balas' suggestion that gender differences may be caused by objectification of female faces, as animacy threshold judgements were not significantly correlated with individual variability in objectification. The effect of gender also cannot be reduced to biases in the doll stimuli used by Balas, or in Experiment One of this paper. It was proposed that the appearance of cosmetics present in female dolls, or the feminine facial features of doll stimuli, such as lightened skin pigmentation and narrow jawlines could have driven the gender effects observed in these experiments. However, in Experiment Two we report comparable gender effects for computer generated inanimate stimuli without such biases. Further research is needed then, to identify the mechanisms behind the influence of gender in animacy perception.

When rating whether a stimulus was able to feel pain, the effect of emotional expression was reversed compared with 'alive' and 'mind' judgements, with lower ratings given to happy faces than neutral. Participants were less likely to attribute the capacity to feel pain to faces that were expressing an emotion incongruent with the experience of pain. In Experiment Two, female faces were judged to be more able to experience pain than male faces, also in contrast with judgements of animacy on the 'alive' and mind' tasks. We propose that this may reflect a gender bias in the perception of others' sensitivity to pain, since both men and women tend to attribute greater pain sensitivity to women than to men (Robinson et al., 2001). While these explanations seem straightforward, the findings suggest that deciding whether or not a face is capable of experiencing pain relies on different stimulus cues to deciding whether the face appears alive or has a mind, and therefore these may reflect two distinct processes. Gray, Gray, and Wegner (2007) propose two core dimensions to mind perception: experience and agency. This account fits with the results of the 'alive' and 'mind' rating tasks. If face was perceived as *experiencing* emotion, then it was associated with appearing more alive, and more likely to have a mind. However, if the capacity to feel pain is to be considered an aspect of the mind, then the experience of emotion should also increase mind perception in this aspect, according to Gray and colleagues. In our stimuli the opposite effect was observed, with emotion expression resulting in reduced perception of the capacity for pain. Further investigation is required to establish whether animacy and the capacity to feel pain reflect separable perceptual judgements.

Finally, in Experiment One, animacy judgements were not significantly affected by whether the stimulus was chromatic or achromatic. This suggests that skin pigmentation is not a vital cue for animacy perception. It should be noted that pigmentation is not the only cue that can be gained from the skin. Texture and shading details are also relevant for other aspects of face processing (Bruce & Langton, 1994; Meinhardt-Injac, Persike, & Meinhardt, 2013), suggesting that these cues may have influenced animacy judgements on the task. However, Looser and Wheatley (2010) show that while animacy ratings for the eyes alone accounted for 75% of the variance of whole-face ratings in their experiment, ratings for skin patches did not account for a significant proportion of this variance. In fact, animacy ratings for the morphed skin patch stimuli increased only by around 10% between 100% inanimate and 100% animate stimuli. This indicates that skin properties do not provide a particularly useful cue to animacy. Overall, our findings support the validity and generalisability of previous studies which have used chromatic images to compare animacy judgements for different stimulus groups (Balas, 2013; Hackel et al., 2014; Looser & Wheatley, 2010; Powers et al., 2014; Swiderska et al., 2012; Wheatley et al., 2011). This provides useful implications for future animacy research, as it appears equally valid to present chromatic or achromatic stimulus faces.

4.2. Perceiver variability factors that contribute to animacy judgments

Prior work has suggested that an individual's desire for social connection may influence the threshold at which animacy is perceived. Powers et al. (2014) report a negative correlation between scores on the NTBS and animacy thresholds, indicating that animacy was perceived more readily by participants with a greater need to belong. This correlation was not replicated in our participant sample in Experiment One. It is worth noting that our sample was larger (104 vs. 30) and showed a wider range of NTBS scores than the sample recruited by Powers and colleagues. It is possible that the extra anonymity provided by an online experiment affected increased the honesty of participants' responses on this self-report measure. Comparison of Internet and lab-based self-report measures indicates that online tests are as reliable as those conducted under controlled conditions (Buchanan & Smith, 1999) and are likely to result in more honest self-disclosure (Joinson, 1999). There was also no significant correlation between self-reported loneliness and animacy perception, as predicted based on the research of Epley et al. (2008). However, the studies conducted by Epley and colleagues focussed on the attribution of anthropomorphic traits to animals and objects, rather than human faces. It appears that the modulating effect of loneliness may be limited to these kinds of stimuli. Overall, while our findings contradict those previously reported, further research is necessary to establish the link between desire for social interaction and the readiness with which animacy is perceived.

In Experiment One, higher levels of externally oriented thinking, a subscale of the TAS measure of alexithymia (Bagby et al., 1994) were associated with increased likelihood of perceiving animacy. However, this effect was not replicated in Experiment Two, which we suggest may be related to the difficulty of making threshold judgements using this stimulus set. The externally oriented thinking subscale is associated with focussing attention on external events rather than inner feelings and emotions (Bagby et al., 1994). It is possible that individuals scoring highly on this subscale may be less influenced by the emotional and social cues in the stimuli, and more influenced by low-level perceptual differences when making animacy judgements. This could result in reduced thresholds for perceiving animacy. The second point to consider is that a lower animacy PSE is also closer to the true point of equality (50%) than a higher PSE. An alternative explanation of the data is therefore that lower animacy thresholds reflect more accurate animacy judgements. Those with an externally oriented cognitive style may therefore be more accurate at animacy perception as a result of relying more on perceptual than social cues. In this study, for instance, a stimulus face that was smiling was no more likely to be animate than a stimulus with neutral expression. Therefore allowing social cues to influence animacy judgements would not improve accuracy. One way of exploring this hypothesis would be to design an animacy task with 'correct' responses. This could involve, for example, a sorting task in which participants must arrange stimulus faces at different animacy levels into the correct answer. This would allow conclusions to be made regarding whether cognitive style and stimulus qualities modulate the accuracy, as well as the likelihood, of animacy perception. At present, however, the link with animacy perception indicates wider differences in face processing in alexithymia than previously thought, which carries important implications for understanding the underlying causal mechanisms.

4.3. Summary

Across two experiments, we establish the effects of both stimulus and individual variability in making animacy judgements. The findings corroborate previous evidence that the threshold at which animacy is perceived along a continuum is influenced by the gender of the stimulus face (Balas, 2013). In addition, we provide novel evidence to suggest that this threshold is also influenced by the emotional expression of the stimulus, with happy faces perceived as being animate at a lower threshold than neutral faces. With regards to individual variability, we find some evidence that an externally oriented thinking style was associated with lower animacy thresholds. Taken as a whole, these findings provide vital guidance for the careful control of stimuli in future facial animacy perception research. They also give rise to broader implications for the study of alexithymia, as well as for the role of cognition, emotion and gender in social perception.

Acknowledgements

MJB was supported by the ESRC [ES/K00882X/1].

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jesp.2017.02.004.

References

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. http://dx.doi.org/10.1016/0022-3999(94)90005-1.

Balas, B. (2013). Biological sex determines whether faces look real. *Visual Cognition*, *21*(6), 766–788. http://dx.doi.org/10.1080/13506285.2013.823138.

Balas, B., & Horski, J. (2012). You can take the eyes out of the doll, but…. *Perception*, *41*(3), 361–364. http://dx.doi.org/10.1068/p7166.

Balas, B., & Koldewyn, K. (2013). Early visual ERP sensitivity to the species and animacy of faces. *Neuropsychologia*, *51*(13), 2876–2881. http://dx.doi.org/10.1016/j.neuropsychologia.2013.09.014.

Balas, B., & Tonsager, C. (2014). Face animacy is not all in the eyes: Evidence from contrast chimeras. *Perception*, *43*(5), 355–367. http://dx.doi.org/10.1068/p7696.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, *133*, 1515–1525. http://dx.doi.org/10.1093/brain/awq060.

Bögels, S. M., & Mansell, W. (2004). Attention processes in the maintenance and treatment of social phobia: Hypervigilance, avoidance and self-focused attention. *Clinical Psychology Review*, *24*(7), 827–856. http://dx.doi.org/10.1016/j.cpr.2004.06.005.

Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception*, *22*(7), 829–840. http://dx.doi.org/10.1068/p220829.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, *23*(7), 803–822. http://dx.doi.org/10.1068/p230803.

Buchanan, T., & Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the world wide web. *British Journal of Psychology*, *90*, 125–144. http://dx.doi.org/10.1348/000712699161189.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, *24*, 723–732. http://dx.doi.org/10.1177/0956797612463582.

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion*, *11*(4), 860–865. http://dx.doi.org/10.1037/a0022758.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*, 581–604. http://dx.doi.org/10.1016/s0149-7634(00)00025-7.

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating social connection through inferential reproduction: Loneliness and perceived agency in gadgets, gods, and greyhounds. *Psychological Science*, *19*, 114–120. http://dx.doi.org/10.1111/j.1467-9280.2008.02056.x.

Franz, M., Popp, K., Schaefer, R., Sitte, W., Schneider, C., Hardt, J., ... Braehler, E. (2008). Alexithymia in the German general population. *Social Psychiatry and Psychiatric Epidemiology*, *43*, 54–62. http://dx.doi.org/10.1007/s00127-007-0265-1.

Frost, P. (1988). Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine*, *32*, 38–58. http://dx.doi.org/10.1353/pbm.1988.0010.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619. http://dx.doi.org/10.1126/science.1134475.

Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., & Berthoz, S. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): Systematic review, unanswered questions and further perspectives. *PloS One*, *7*(8), e42429. http://dx.doi.org/10.1371/journal.pone.0042429.

Hackel, L. M., Looser, C. E., & Van Bavel, J. J. (2014). Group membership alters the threshold for mind perception: The role of social identity, collective identification, and intergroup threat. *Journal of Experimental Social Psychology*, *52*, 15–23. http://dx.doi.org/10.1016/j.jesp.2013.12.001.

Heinrichs, N., & Hofmann, S. G. (2001). Information processing in social phobia: A critical review. *Clinical Psychology Review*, *21*(5), 751–770. http://dx.doi.org/10.1016/s0272-7358(00)00067-2.

Joinson, A. (1999). Social desirability, anonymity and internet-based questionnaires. *Behavior Research Methods, Instruments and Computers*, *31*(3), 433–438. http://dx.doi.org/10.3758/bf03200723.

Jongen, S., Axmacher, N., Kremers, N. A., Hoffmann, H., Limbrecht-Ecklundt, K., Traue, H. C., & Kessler, H. (2014). An investigation of facial emotion recognition impairments in alexithymia and its neural correlates. *Behavioural Brain Research*, *271*, 129–139. http://dx.doi.org/10.1016/j.bbr.2014.05.069.

Keltner, D., & Kring, A. M. (1998). Emotion, social function and psychopathology. *Review of General Psychology*, *2*(3), 320–342. http://dx.doi.org/10.1037//1089-2680.2.3.320.

Lane, R. D., Sechrest, L., Reidel, R., Weldon, V., Kaszniak, A., & Schwartz, G. E. (1996). Impaired verbal and nonverbal emotion recognition in alexithymia. *Psychosomatic Medicine*, *58*, 203–210. http://dx.doi.org/10.1097/00006842-199605000-00002.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition & Emotion*, *24*(8), 1377–1388. http://dx.doi.org/10.1080/02699930903485076.

Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59. http://dx.doi.org/10.1016/s1364-6613(99)01436-9.

Leary, M. R., Kelly, K. M., Cottrell, C. A., & Schreindorfer, L. S. (2013). Construct validity of the need to belong scale: Mapping the nomological network. *Journal of Personality Assessment*, *95*, 610–624. http://dx.doi.org/10.1080/00223891.2013.819511.

Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy. *Psychological Science*, *21*(12), 1854–1862. http://dx.doi.org/10.1177/0956797610388044.

Looser, C. E., Guntupalli, J. S., & Wheatley, T. (2013). Multivoxel patterns in face-sensitive temporal regions reveal an encoding schema based on detecting life in a face. *Social Cognitive and Affective Neuroscience*, *8*(7), 799–805. http://dx.doi.org/10.1093/scan/nss078.

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2013). Holistic face processing is induced by shape and texture. *Perception*, *42*(7), 716–732. http://dx.doi.org/10.1068/p7462.

Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., ... Komaki, G. (2006). Impaired self-awareness and theory of mind: An fMRI study of mentalizing in alexithymia. *NeuroImage*, *32*(3), 1472–1482. http://dx.doi.org/10.1016/j.neuroimage.2006.04.186.

Noll, S. M. (1996). *The relationship between sexual objectification and disordered eating. Correlational and experimental tests of body shame as a mediator.* Unpublished doctoral dissertation. Durham, NC.: Duke University.

Noll, S. M., & Fredrickson, B. L. (1998). A meditational model linking self-objectification, body shame and disordered eating. *Psychology of Women Quarterly*, *22*, 623–636. http://dx.doi.org/10.1111/j.1471-6402.1998.tb00181.x.

Pandey, R., & Mandal, M. K. (1997). Processing of facial expressions of emotion and alexithymia. *British Journal of Clinical Psychology*, *36*(4), 631–633. http://dx.doi.org/10.1111/j.2044-8260.1997.tb01269.x.

Parker, J. D. A., Taylor, G. J., & Bagby, R. M. (1993). Alexithymia and the recognition of facial expressions of emotion. *Psychotherapy and Psychosomatics*, *59*, 197–202. http://dx.doi.org/10.1159/000288664.

Parker, J. D. A., Taylor, G. J., & Bagby, R. M. (2001). The relationship between emotional intelligence and alexithymia. *Personality and Individual Differences*, *30*, 107–115. http://dx.doi.org/10.1016/s0191-8869(00)00014-3.

Peters, L., Sunderland, M., Andrews, G., Rapee, R. M., & Mattick, R. P. (2012). Development of a short form Social Interaction Anxiety (SIAS) and Social Phobia Scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. *Psychological Assessment*, *24*(1), 66–76. http://dx.doi.org/10.1037/a0024544.

Powers, K. E., Worsham, A. L., Freeman, J. B., Wheatley, T., & Heatherton, T. F. (2014). Social connection modulates perceptions of animacy. *Psychological Science*, *25*(10), 1943–1948. http://dx.doi.org/10.1177/0956797614547706.

Robinson, M. E., Riley, J. L., Myers, C. D., Papas, R. K., Wise, E. A., Waxenberg, L. B., & Fillingim, R. B. (2001). Gender role expectations of pain: Relationship to sex differences in pain. *The Journal of Pain*, *2*(5), 251–257. http://dx.doi.org/10.1054/jpai.2001.24551.

Russell, D. W. (1996). UCLA loneliness scale (version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, *66*, 20–40. http://dx.doi.org/10.1207/s15327752jpa6601_2.

Strelan, P., & Hargreaves, D. (2005). Women who objectify other women: The vicious circle of objectification? *Sex Roles*, *52*, 707–711. http://dx.doi.org/10.1007/s11199-005-3737-3.

Swiderska, A., Krumhuber, E., & Kappas, A. (2012, September). Perception of animacy in Caucasian and Indian faces. *Paper presented at ACM Third International Symposium on Facial Analysis and Animation (FAA), Vienna.* http://dx.doi.org/10.1145/2491599.2491608.

Wheatley, T., Weinberg, A., Looser, C., Moran, T., & Hajcak, G. (2011). Mind perception: Real but not artificial faces sustain neural activity beyond the N170/VPP. *PloS One*, *6*(3), e17960. http://dx.doi.org/10.1371/journal.pone.0017960.

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the east and west? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, *43*(2), 303–311. http://dx.doi.org/10.1016/j.jesp.2006.02.004.