Q-cgi: new techniques to assess variation in perception applied to facial attractiveness

D. Michael Burt[1,*], **Robert W. Kentridge**[1], **James M. M. Good**[1], **David I. Perrett**[2], **Bernard P. Tiddeman**[3] **and Lynda G. Boothroyd**[1]

[1]*Department of Psychology, University of Durham, South Road, Durham DH1 3LE, UK*
[2]*School of Psychology, and* [3]*School of Computer Science, University of St Andrews, Fife KY16 9AJ, UK*

We present novel methods for assessing variation in the perception of subjective cues based on a fusion of Q-methodology with computer graphics techniques. Participants first Q-sort face stimuli based upon a subjective quality; a randomization-based statistic is then calculated to test whether groups of participants differ in their perception. Computer graphics are then used to extract and illustrate the differences in the manner which participants sorted so that the differences can be quantified. As a demonstration, the technique is applied to investigate the effects of prospective relationship duration and of sexual restrictiveness on the characteristics which participants find attractive in photographs of opposite-sex faces. Results show that in a naturally varying set of faces, female participants prefer facial cues related to masculinity for short-term relationships, whereas characteristics related to positive personality attributes are preferred for long-term relationships. For short-term relationships, male participants appear to prefer more feminine, youthful faces. Preferences of individuals with less restricted sexual strategy paralleled short-term preferences in that more feminine female faces and more masculine male faces were preferred.

Keywords: facial attractiveness; relationships; masculinity; computer graphics; Q-methodology

1. INTRODUCTION

We present a new methodology for assessing subjective variation between groups of participants. Variation in participants' perception is recorded using Q-sort, a component of Q-methodology (Stephenson 1953), in which participants allocate their judgements of items according to a fixed distribution (figure 2). The probability that two groups of participants differ in their manner of sorting, and thus their perception of the items more than would be expected by chance, is then assessed using a novel technique based on randomization. This randomization technique can be applied to any experiment in which participants sort items according to a fixed symmetrical distribution, including ranking data. Computer graphics are then used to enable the differences between groups of Q-sorts to be visualized. Here, we illustrate the technique by an investigation of variation in the perception of facial attractiveness to examine the differences in what is perceived attractive by sexually restricted individuals when compared with less restricted individuals and comparing individuals' perceptions of attractiveness in long- and short-term relationship scenarios.

The manner in which the current investigation was conducted is different from many previous studies of facial attractiveness. Rather than setting out with an experiment focused on a predetermined attribute, such as health (Jones *et al.* 2005), masculinity (Perrett *et al.* 1998; Penton-Voak *et al.* 1999; Rhodes *et al.* 2000; Johnston *et al.* 2001; Little *et al.* 2002) or symmetry (Rhodes *et al.* 1998;

Perrett *et al.* 1999), participants were asked to judge the attractiveness of a sample of facial photographs of real people. The computer graphics analysis then enabled commonalities in participants' preference to emerge from the data. Subject to the limitations of the computer graphics technique, this method captures the complex, multidimensional factors that mediate agreement in participants' perception rather than focusing on predetermined factors. Thus, if factors such as masculinity and personality are important in the real world, as suggested by previous highly controlled studies, then these factors would be expected to emerge from the analysis of the Q-sort data.

To capture variation in perception of attractiveness, participants were presented with a set of 30 opposite-sex faces to Q-sort, first based on how attractive they felt each of the faces would be for a long-term relationship and then based on how attractive each of the faces would be for a short-term relationship. As an illustration of the typical use of Q-sort and Q-methodology, Brown (1996) investigated the type of attention a patient received from different carers. The patient evaluated a set of statements based on how well each statement reflected the care given by each carer. Instead of using a Likert-type scale, a scale in which the patient had to allocate ratings to the statements according to a fixed quasinormal distribution, similar to that shown in figure 1, was used. The patient could select only one item as best describing each of their carers. The study of facial attractiveness presented here was carried out in the same manner, thus, for example, in the long-term relationship context, each participant can give only one of the 30 faces the highest rating, whereas six of the faces must receive the median rating. Typically, in Q-methodology, the data are then subjected to factor
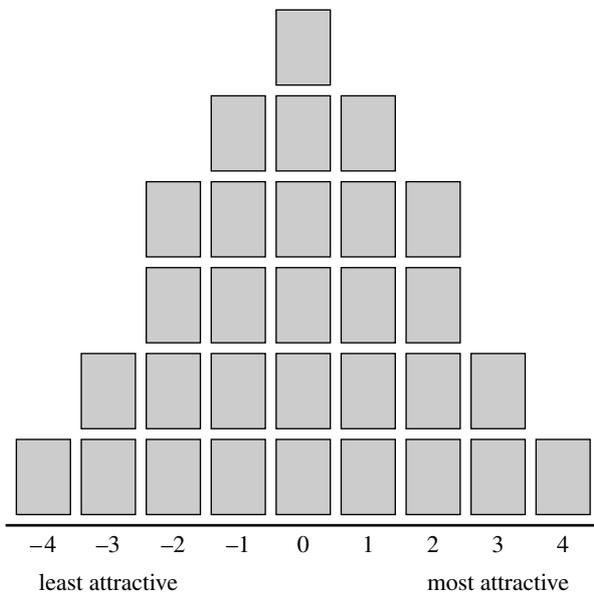
Figure 1. An illustration of the grid participants used for Q-sorting the images. The grid fixes the distribution of rating that participants are able to give.

analysis to extract the patterns of differences and similarities between sorts. In Brown's study, commonalities focusing on professional matters were found in the care given by the surgeon, nurse A, and the care given during the process of the hospitalization. This was distinguished from the more empowering care given by nurse C and the compassion-focused care the patient received from his mother.

Although Q-methodology is not limited to the use of textual items, the textual items are most often used as they are easily described and summarized as factors. However, as faces defy concise description, a computer graphics technique was used to summarize facial attributes perceived differently by groups of participants. Similar computer image warping and prototyping procedures have been used previously to extract commonalities in groups of face images (Langlois & Roggman 1990; Benson & Perrett 1991; Burt & Perrett 1995; Perrett *et al.* 1998; Tiddeman *et al.* 2001). However, instead of an average-based technique, a technique based on regressions of values pertaining to the images was used here. The use of regression is more sensitive as it uses the whole rating scale and all of the image stimuli within a set, rather than, for example, comparing a composite of image stimuli rated high on a scale with a composite of image stimuli rated low. However, it is important to note that the method is linear and will be insensitive to nonlinear and sparsely distributed stimulus variations. To differentiate the images produced in the current study from averages or prototypes produced by other techniques, these images will be referred to as Q-cgi images, reflecting the combination of Q and computer graphics methods.

2. MATERIAL AND METHODS

(a) Participants
A total of 50 participants (26 males) aged between 18 and 22 years ($M = 20.2$) took part in the experiment. One of the female participants reported that she was homosexual and was removed from the analysis as the experiment was to examine the opposite-sex preferences. The second part of this experiment involved 45 of the same participants (23 males).

(b) Stimuli for Q-sorting
The faces of 30 male and 30 female students, who were photographed under standard conditions (Perrett *et al.* 1998) while being asked to smile, were selected from a large database. Faces were selected on the basis of being white (as variation in skin colour was not being studied) and lacking facial hair and highly distinctive features like facial jewellery. Features outside the face, including the hair and ears, were obscured and each face image was printed in photographic quality on 4×5 cm cards.

(c) Procedure
After completing a questionnaire which included the socio-sexual orientation inventory (SOI; Simpson & Gangestad 1991), participants were provided with written definitions for long- and short-term relationships (Perrett *et al.* 2002). Participants completed two Q-sorts for the experiment. In the first, participants were asked to sort opposite-sex faces based on how attractive they perceived the faces to be for a long-term relationship. When finished, participants were asked to Q-sort the same faces based on how attractive the faces were for a short-term relationship. To help participants Q-sort, a Q-sort grid illustrated in figure 1 was constructed, which ranged from least attractive (rating of -4) to most attractive face (rating of $+4$). After completing the Q-sorts, a period of at least four weeks was allowed to elapse before a subset of the participants were presented with the Q-cgi face pairs and asked to make forced choice personality judgements.

(d) Q-cgi face preparation
Regression was used to relate the parameters describing the colour and shape of Q-sorted face images (e.g. the redness of the pixel on the tip of the nose, the height of the right lip corner etc) to a value k, the difference in average rating received by the face image stimulus under two conditions of Q-sorting. The method is illustrated in figure 2 when k is the difference in ratings given to face stimuli in long-term relationship as opposed to short-term relationship scenarios. Using the calculated regression coefficients, an estimate of each of the parameters describing the face can then be computed for any level of k and these can be reassembled to form a face image.

Processing of the male and female face images was performed separately. For each image, the positions of 208 feature points were used to demarcate the position and shape of the facial features (Benson & Perrett 1991) and calculate the average face shape. Image warping was then used to bring each image into correspondence with the average face shape. Parameters representing the corresponding red, green and blue pixel values in each of the warped face images and the corresponding x and y feature point values of each of the face shapes were then separately regressed with the k value.

To create the face image corresponding to a given value of k, the computed regression coefficients were used to calculate the x and y feature point values denoting the face shape and the red, green and blue pixel values denoting the image. The final Q-cgi image was made by warping the resultant image from the average face shape to the calculated face shape. Features outside the face were then obscured.

The above method was used to compose (i) male and female Q-cgi images with nominal long-term preference
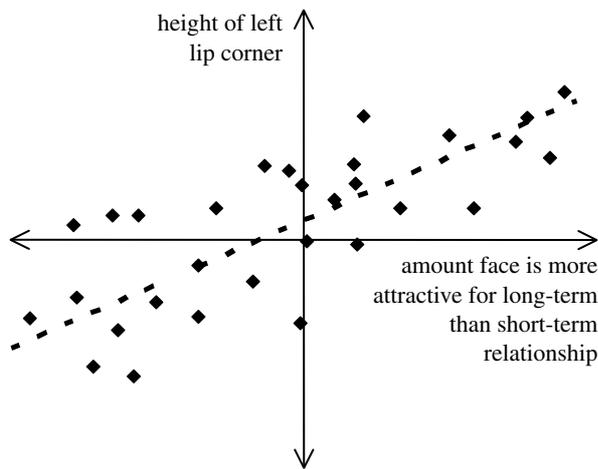
Figure 2. Illustration of the relationship between the height of the left lip corner in each face stimulus and the amount that participants preferred the face for a long-term relationship than for a short-term relationship. Using the regression line, a value for the height of the left lip corner can be calculated for any level of preference. The computer graphics analysis works by calculating (i) this linear regression separately for each of the *x* and *y* coordinates of each of the feature points to produce a face shape for a given value of preference and (ii) the same analysis separately on each of the *r*, *g* and *b* values of each of the pixels in each of the face images to give the colour values. (Data for illustration purposes only.)

scores of 1 and −1 and (ii) male and female Q-cgi images with nominal restricted preference scores of 1 and −1 by subtracting the average Q-ratings given to each face stimulus by individuals who scored above and below the mean behavioural SOI score. Behavioural SOI score was calculated following the method set out by Simpson (1998) with the exception that the attitudinal component of the questionnaire was disregarded.

(e) *Statistical test for a difference in the manner of sorting*

As the distribution from which each Q-sort is made is fixed and known, it is possible to assess the probability that differences in the patterns of Q-sorting under two conditions are greater than would be expected by chance (and are instead due to different patterns of Q-sorting under the two conditions). As an index of how different the pattern of Q-sorting is under two conditions, we used the average unsigned difference in rating given to each item under the two conditions (with *M* and *N* being the numbers of Q-sorts completed in each condition). While the sampling distribution of this statistic can be calculated when the number of Q-sorts in each condition is the same, there is no straightforward analytical means of estimating the parameters of the sampling distribution when the number of Q-sorts in each condition differs. We, therefore, adopted a randomization approach to test for differences between groups.

Randomization tests are typically carried out using a computer to simulate repeatedly rerunning the experiment by randomly redistributing the observed experimental data (1000 at minimum; Manly 1991) to produce an estimate of the sampling distribution for the statistic of interest. We also simulated random reruns of our experiment. However, in contrast to the conditions facing most experimenters, we know the distribution of the entire population from which each of our observations is drawn (owing to the fixed

distribution imposed by Q-sort). Hence, instead of generating reruns from a single set of observations, we can generate each Q-sort in our reruns by sampling at random from the population distribution. To obtain a sampling distribution against which our experimentally derived index of the magnitude of difference in the pattern of Q-sorting between conditions can be tested, the experiment was simulated 10 000 times with *M* and *N* Q-sorts in each condition. (An EXCEL file with the commented VISUAL BASIC macro used to perform this calculation is included in the electronic supplementary material, which can be used for the reader's data as well as the example data given.) An estimate of the *p*-value can then be determined to assess the likelihood that the difference between groups was due to chance.

Our procedure involves simulating participants Q-sorting randomly and will produce an overestimate of *p* when there is high agreement between participants. Thus, high levels of agreement will render the test more conservative. However, as illustrated in the electronic supplementary material, the inflation of *p* may be particularly extreme when the same participants are compared under two similar conditions of Q-sorting, such as Q-sorting faces for attractiveness under long- and short-term scenarios. A correction for a between-participants design can be used, in which data are analysed from only one condition per participant. The value of the average unsigned difference will vary depending on which participants are allocated to the two conditions. To provide a general index of the average unsigned difference for the comparison, the allocation must be performed many times to return an average of average unsigned differences with half of the participants allocated to each condition. A function for this calculation is provided in the electronic supplementary material.

(f) *Presentation of Q-cgi stimuli*
Participants were presented with pairs of opposite-sex Q-cgi image stimuli characterizing the differences in (i) long- and short-term preference and (ii) preference between individuals with a restricted and unrestricted sexual strategy in five randomly ordered blocks. In each block, all participants judged which face had more of a personality-related trait (health, youthfulness, masculinity, kindness, trustworthiness or good parenting skills) using an 8-point scale. The scale enabled participants to indicate simultaneously the face which they felt possessed more of the trait and their level of confidence (1–4 points) that the pair of faces differed on the trait.

3. RESULTS AND DISCUSSION
(a) *Effect of long- and short-term relationship context*
Initial analysis using the randomization technique revealed no effect of relationship term on sorting of faces by either male (average unsigned difference (AUD) = 0.22, *p* < 0.26) or female participants (AUD = 0.26, *p* < 0.12). However, as noted in §2 and demonstrated in the electronic supplementary material, high levels of agreement make the statistic more conservative, an effect which can be acute when the same participant performs similar judgements in two conditions. To correct for this effect, the sampling distribution of the average unsigned difference was estimated between participants. After performing this correction, there were significant
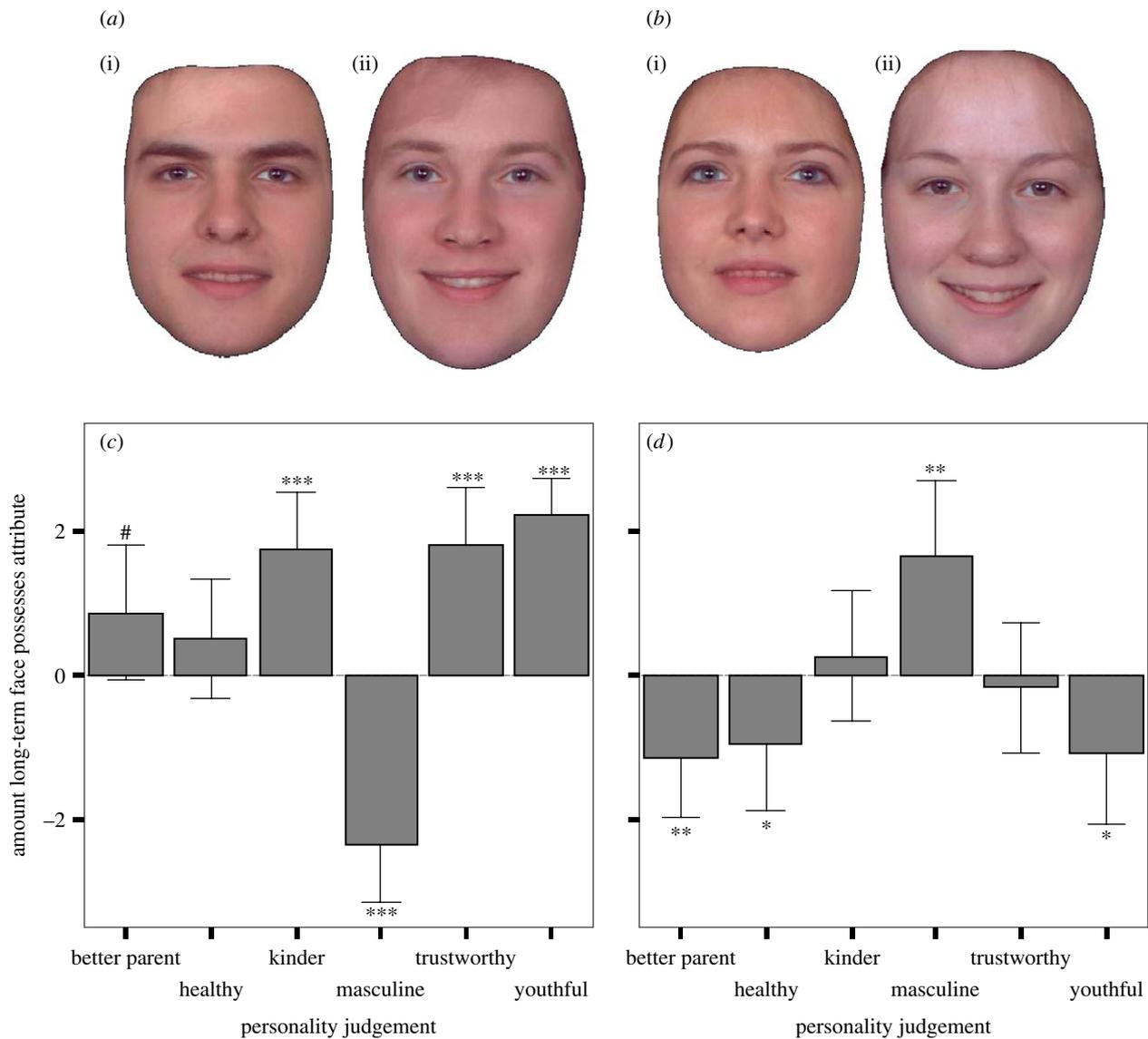
Figure 3. Effects of level of relationship term on perceptions of facial attractiveness. Pairs of Q-cgi faces (*a*, male faces; *b*, female faces) constructed based on attractiveness judgements made by participants in (i) short- and (ii) long-term relationship scenarios. (*a*(i),*b*(i)) is based on the characteristics of faces found more attractive for a short-term relationship and (*a*(ii),*b*(ii)) is based on the characteristics found more attractive for a long-term relationship. (*c*,*d*) The graphs represent participant personality attributions to the faces given immediately above. Scores that are higher than zero reflect participants ascribing a personality attribute to the face image characterizing attractiveness preferences for a long-term relationship (*a*(ii), *b*(ii)) rather than to the other image. Symbols represent level of significance: #$p < 0.1$, *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$ with 95% CIs.

differences in the manner that both male (AUD$=0.45$, $p < 0.001$) and female (AUD$=0.51$, $p < 0.001$) participants sort under long- and short-term relationship conditions. To enable the differences in the manner of sorting under long- and short-term relationship conditions to be visualized, Q-cgi face images were produced and are shown in figure 3*a*,*b*, which summarize participants' perception of the Q-cgi images based on different personality-related traits.

Results were in line with previous studies (Penton-Voak *et al.* 1999; Little *et al.* 2002; Scarbrough & Johnston 2005): the image characterizing the male faces that were preferred for short-term relationships appears to be more masculine than the image characterizing the male faces that were preferred for long-term relationships. This 'long-term' male face was perceived as being more feminine, kinder, more trustworthy and more youthful although there was a trend for participants to consider the

'short-term' male face a better parent. These characteristics have been previously found to relate to lower levels of masculinity in male faces (Perrett *et al.* 1998). Female participants' preference for more masculine faces in short-term relationships is consistent with a bias for selecting cues to good genes in short-term relationship scenarios (Little *et al.* 2002). However, neither face was perceived as significantly healthier.

As can be seen from figure 3*b*, the female face image characterizing faces preferred by males for short-term relationships appeared more feminine than that characterized preferences for long-term relationships. The feminine 'short-term' female face image was also attributed to other positive personality-related traits of being a better parent, being healthier and more youthful, a pattern consistent with cues to fertility being more important to male participants for short- than for long-term relationships.
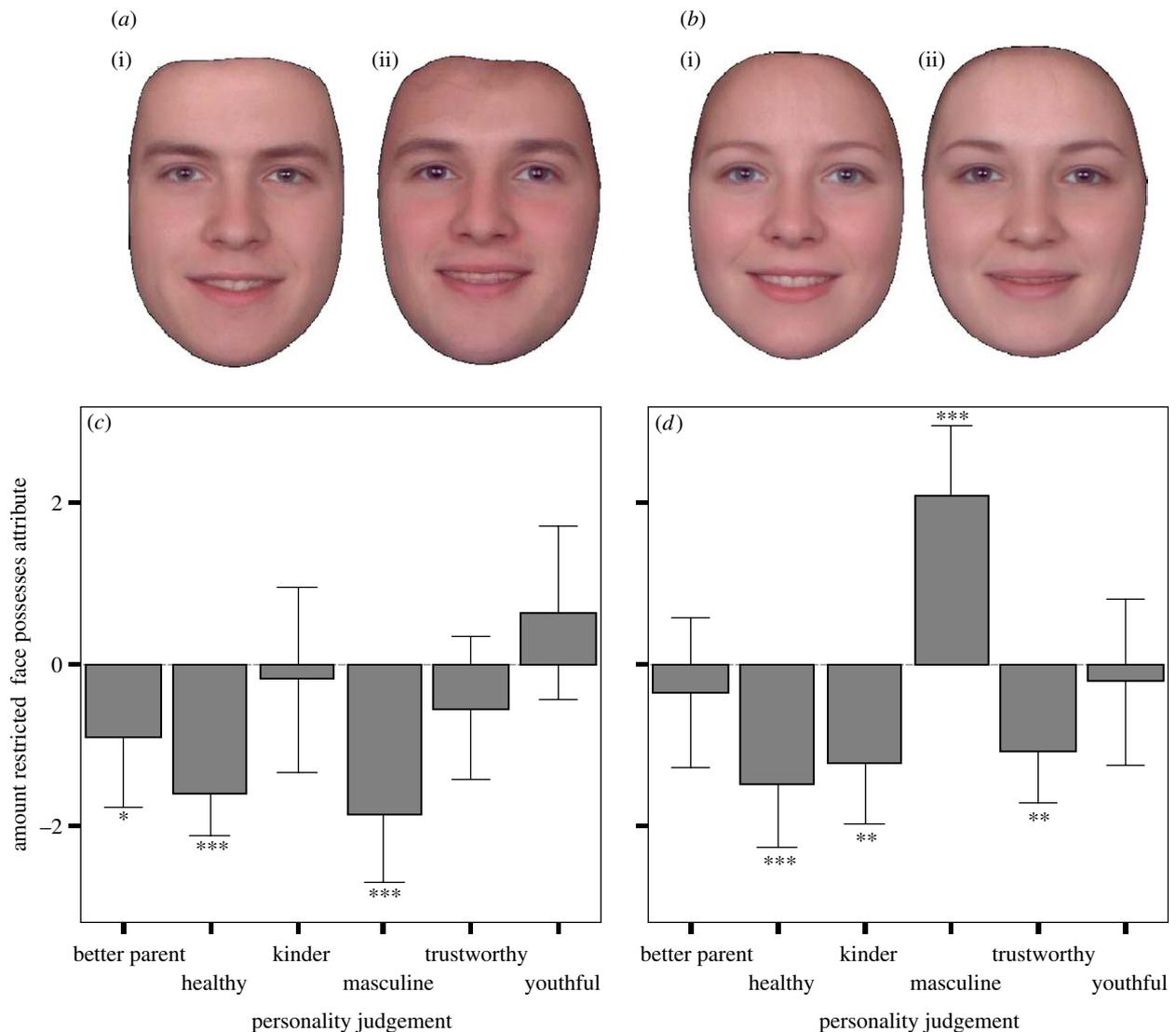
Figure 4. Effects of restrictedness in sexual strategy on perceptions of facial attractiveness. Pairs of Q-cgi faces (*a*, male faces; *b*, female faces) constructed based on the attractiveness judgements of (i) unrestricted and (ii) restricted participants. (*a*(i),*b*(i)) is based on the characteristics of faces found more attractive by the unrestricted individuals and (*a*(ii),*b*(ii)) is based on the characteristics found more attractive by restricted individuals. (*c*,*d*) The graphs represent participant personality attributions to the faces immediately above. Scores that are higher than zero reflect participants attributing more of a personality attribute to the face image characterizing the attractiveness preferences of restricted participants (*a*(ii),*b*(ii)) than to the other image. Symbols represent the level of significance: *$p<0.05$, **$p<0.01$ and ***$p<0.001$ with 95% CIs.

Participants were not asked to assess facial expression. However, one immediately noticeable characteristic of both the male and female Q-cgi faces characterizing long-term preferences is that they appear to be smiling more than the faces characterizing short-term relationship preferences. This pattern may be indicative of the greater importance of a positive personality for a long-term relationship for both sexes.

(b) *Effects of sexual restrictedness on perception of attractiveness in a long-term relationship scenario*

Participants were assigned to sexually restricted or unrestricted groups based on whether their behavioural SOI (scored from the number of past and predicted future sexual partners and extra-pair fantasies) was higher or lower than the average score for the sample of all participants. Unrestricted female participants ($n=8$) showed a significantly different pattern of sorting faces for attractiveness (AUD$=0.48$, $p<0.01$) than restricted female participants ($n=15$). Unrestricted male

participants ($n=15$) showed a significantly different pattern of sorting faces for long-term attractiveness (AUD$=0.41$, $p<0.01$) than restricted male participants ($n=11$). To enable these differences to be visualized, Q-cgi face images were produced, which characterize the differences in what is found attractive by the restricted and unrestricted participants.

Figure 4 presents the Q-cgi images illustrating the effect of participants' sexual restrictedness on their perceptions of long-term attractiveness. The less restricted female participants, in comparison with the more restricted female participants, preferred male faces that were characterized as being more masculine, healthy and a better parent. Less restricted male participants preferred faces that were characterized as being more feminine, healthier, kinder and more trustworthy than those preferred by the more restricted males. The preferences expressed by less restricted participants for cues to possible markers of testosterone in males and oestrogen in females are in keeping with the notion that cues to good

genes and fertility may be particularly important to participants pursuing a less restricted strategy (Little *et al.* 2002). However, the male face that was composed based on the preferences of unrestricted women was also perceived to be a better parent and, likewise, the corresponding female face was perceived as being kinder and more trustworthy. In the case of the male face, attributions of being a better parent may be due to the relationship between masculinity and maturity with more mature males being seen as better fathers. Similarly, higher levels of facial femininity in the female face may be linked to both appearing to be kinder and more trustworthy. Alternatively, less restricted individuals may be less concerned with mate retention, possibly, because these individuals may be more attractive and find it easier to retain a mate, leading them to select faces that are generally more attractive.

4. CONCLUSIONS

We have demonstrated the functionality of a new technique, Q-cgi, to assess variation in preference and have applied the technique to investigate difference in participant preferences for faces in short- and long-term relationship scenarios and between participants who are more or less sexually restricted. Through comparison of our findings with the previous literature, it is concluded that characteristics manipulated by previous studies in a highly controlled manner (e.g. Perrett *et al.* 1998; Rhodes *et al.* 2000; Johnston *et al.* 2001; Penton-Voak *et al.* 2003) are important enough to influence the manner in which a naturalistic set of photographs of real faces are perceived compared with the multitude of other facial cues available. Furthermore, although the similarities between the findings presented here and in previous studies suggest that the findings presented here are generalizable, studies using other stimuli or presenting stimuli to different participants may find different results.

REFERENCES

Benson, P. J. & Perrett, D. I. 1991 Synthesising continuous-tone caricatures. *Image Vision Comput.* **9**, 123–129. (doi:10.1016/0262-8856(91)90022-H)

Brown, S. R. 1996 Q Methodology and qualitative research. *Qual. Health Res.* **6**, 561–567. (doi:10.1177/104973239 600600408)

Burt, D. M. & Perrett, D. I. 1995 Perception of age in adult Caucasian male faces—computer graphic manipulation of shape and colour information. *Proc. R. Soc. B* **259**, 137–143. (doi:10.1098/rspb.1995.0021)

Johnston, V. S., Hagel, R., Franklin, M., Fink, B. & Grammer, K. 2001 Male facial attractiveness—evidence for hormone-mediated adaptive design. *Evol. Hum. Behav.* **22**, 251–267. (doi:10.1016/S1090-5138(01)00066-6)

Jones, B. C. *et al.* 2005 Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proc. R. Soc. B* **272**, 347–354. (doi:10.1098/rspb. 2004.2962)

Langlois, J. H. & Roggman, L. A. 1990 Attractive faces are only average. *Psychol. Sci.* **1**, 115–121. (doi:10.1111/ j.1467-9280.1990.tb00079.x)

Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M. & Perrett, D. I. 2002 Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. R. Soc. B* **269**, 1095–1100. (doi:10.1098/rspb.2002.1984)

Manly, B. F. J. 1991 *Randomization and Monte Carlo methods in biology.* London, UK: Chapman & Hall.

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K. & Minamisawa, R. 1999 Menstrual cycle alters face preference. *Nature* **399**, 741–742. (doi:10.1038/21557)

Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P. & Perrett, D. I. 2003 Female condition influences preferences for sexual dimorphism in faces of male *Homo sapiens*. *J. Comp. Psychol.* **117**, 264–271. (doi:10.1037/0735-7036.117.3.264)

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D. & Akamatsu, S. 1998 Effects of sexual dimorphism on facial attractiveness. *Nature* **394**, 884–887. (doi:10.1038/29772)

Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A. & Edwards, R. 1999 Symmetry and human facial attractiveness. *Evol. Hum. Behav.* **20**, 295–307. (doi:10.1016/S1090-5138(99)00014-8)

Perrett, D. I., Penton-Voak, I. S., Little, A. C., Tiddeman, B. P., Burt, D. M., Schmidt, N., Oxley, R., Kinlock, N. & Barrett, L. 2002 Facial attractiveness judgements reflect learning of parental age characteristics. *Proc. R. Soc. B* **269**, 873–880. (doi:10.1098/rspb.2002.1971)

Rhodes, G., Proffitt, F., Grady, J. M. & Sumich, A. 1998 Facial symmetry and the perception of beauty. *Psychon. Bull. Rev.* **5**, 659–669.

Rhodes, G., Hickford, C. & Jeffery, L. 2000 Sex-typicality and attractiveness: are supermale and superfemale faces super-attractive. *Br. J. Psychol.* **91**, 125–140. (doi:10.1348/ 000712600161718)

Scarbrough, P. & Johnston, V. 2005 Individual differences in women's facial preferences as a function of digit ratio and mental rotation ability. *Evol. Hum. Behav.* **26**, 509–526. (doi:10.1016/j.evolhumbehav.2005.03.002)

Simpson, J. A. 1998 Sociosexual orientation inventory. In *Handbook of sexuality-related measures* (eds C. M. Davis, W. L. Yarber, R. Bauserman, R. Schreer & S. L. Davis), pp. 565–567. Thousand Oaks, CA: Sage.

Simpson, J. A. & Gangestad, S. W. 1991 Individual differences in sociosexuality: evidence for convergent and discriminant validity. *J. Pers. Soc. Psychol.* **60**, 870–883. (doi:10.1037/0022-3514.60.6.870)

Stephenson, W. 1953 *The study of behavior: Q-technique and its methodology.* Chicago, IL: University of Chicago Press.

Tiddeman, B., Burt, M. & Perrett, D. 2001 Computer graphics in facial perception research. *IEEE Comput. Graph. Appl.* **21**, 42–50. (doi:10.1109/38.946630)