Static and Dynamic Facial Images Cue Similar Attractiveness Judgements

S. Craig Roberts*, Tamsin K. Saxton*, Alice K. Murray*, Robert P. Burriss*, Hannah M. Rowland* & Anthony C. Little†

* School of Biological Sciences, University of Liverpool, Liverpool, UK

† School of Psychology, University of Stirling, Stirling, UK

Abstract

Approaches to the study of human mate preferences commonly involve judgements of facial photographs and assume that these judgements provide a reasonable reflection of how individuals would be perceived in real encounters. However, three recent studies have each reported non-significant correlations between judgements using photos (static images) and those using videos (dynamic images). These results have led previous authors to conclude that static and dynamic faces are judged according to different evaluative standards and that this may call into question the validity of findings from experiments using static images. However, the extent of the discrepancy in judgements between image formats remains unknown, and may be influenced by different experimental designs. Here, we tested the effects of several experimental design factors on the strength of correlations between image presentation formats. Using both male and female targets, we compared observed static–dynamic judgement correlations when (1) judgements were made by the same or different raters, or (2) by raters of the same- or opposite-sex to the targets, and (3) when dynamic stimuli were collected under different contextual scenarios. For (1) and (2), we also measured correlations when order of presentation of static and dynamic stimuli was alternated. Our results suggest that each design feature has independent effects on the strength of static–dynamic correlations. Correlations were stronger when static and dynamic stimuli were rated by the same raters. They were weakest for judgements of males by females, when based on seeing photos before videos. This interaction with sex is consistent with previous studies, indicating that females are especially responsive to male dynamic cues. However, in contrast to previous findings and in all cases, static–dynamic correlations were strongly and significantly positive, indicating that judgments based on static images provide an accurate representation of someone's attractiveness during prolonged encounters.

Introduction

There is enormous interest in the biological underpinning of human mate choice judgements (reviews in Grammer et al. 2003; Gangestad & Scheyd 2005; Roberts & Little 2008). While research in this area addresses many different facets of mate preference, by far the most attention focuses on facial attractiveness (Perrett et al. 1994; Penton-Voak & Perrett 2000; Rhodes 2006). Numerous studies have

revealed that different aspects of facial phenotype influence preferences, ranging from robust and universal indicators of mate quality such as sexual dimorphism, symmetry, health and genetic heterozygosity (Perrett et al. 1998; Thornhill & Gangestad 1999; Penton-Voak et al. 2001; Jones et al. 2005; Roberts et al. 2005a) through to more subtle, individual effects associated with menstrual cycle variation in women, self-assessed mate quality, physical resemblance or genetic complementarity (Penton-Voak et al. 1999; Little et al. 2001; Roberts et al. 2004, 2005b; DeBruine et al. 2005). These and many other facial preference studies commonly involve judgements of facial photographs. Because researchers have often been interested in structural aspects of the face and wish to control for personality cues, expression or gaze, target participants are typically photographed looking towards the camera with a neutral expression.

More recently, there has been growing recognition that neutral photographs (static images) portray only part of the information that is used when forming mate preference judgements. While structural aspects of faces are likely to explain a large proportion of the variance in real-life judgements, a significant proportion will also be accounted for by variation in expressiveness, gaze, perceived personality and other cues which are available during actual encounters and in dynamic images such as videos (Grammer 1990; Grammer et al. 2000; Gangestad et al. 2004; Morrison et al. 2007; Conway et al.

2008; Penton-Voak & Chang 2008). Indeed, the use of videos as a more ethological, and ecologically valid, approach has been recognised as an important and emerging dimension to the study of human mate preferences (Gangestad & Scheyd 2005; Penton-Voak & Chang 2008; Roberts 2008).

Against this background, interpreting the relative importance of structural and movement cues is a critical issue. In other words, it is important to estimate the extent to which judgements made under static and dynamic presentations either do or do not reflect each other. In particular, if judgements of static images do not predict judgements using dynamic images, this would call into question the validity of a significant amount of existing research using photographs alone and radically challenge the foundations on which the research field is based. However, there have been surprisingly few direct tests of the relationship between judgements based on static and dynamic images (Penton-Voak & Chang 2008) and the results of the studies conducted so far are mixed (see Table 1). Rubenstein (2005) found that judgements of female images under the two conditions were not strongly, and not significantly, correlated. He concluded that static and dynamic faces are judged according to different evaluative standards, in which the latter are particularly influenced by emotional expression. In contrast, Lander (2008) found significant static–dynamic correlations for judgements of female images, but not for images of males. Extending Rubenstein's work, Penton-Voak & Chang

Table 1: Summary of the methodology and results of studies investigating correlations between attractiveness ratings when images are presented in static and dynamic formats

Author(s)	Stimulus sex	Rater sex	Context	Design	Sound	Correlation
Rubenstein (2005)	F	MF	Cue	BS	Yes?	ns
Lander (2008)	M	M	Cue	BS	No	ns
	M	F	Cue	BS	No	ns
	F	M	Cue	BS	No	+
	F	F	Cue	BS	No	+
Penton-Voak & Chang (2008)	M	MF	Cue	BS	No	ns
	M	MF	Holiday	BS	No	ns
	F	MF	Cue	BS	No	+
	F	MF	Holiday	BS	No	+
Roberts et al. (2009)	M	F	Introd.	WS	No	+
Current study	M	M,F,MF	Introd.	BS,WS	No	+
	F	M,F,MF	Introd.	BS,WS	No	+
	M	F	Holiday	WS	No	+

Stimulus sex: M, male; F, female. Rater sex: MF, all raters; the term M,F,MF indicates separate analyses using either M raters, F raters or MF raters. Context denotes the task given to stimuli when recording dynamic images (Cue, cue card; Holiday, description of holiday plans; Introd., introductory video to member of opposite-sex). Design denotes whether the same or different raters judged static and dynamic images (BS, between subjects; WS, within subjects). Correlation denotes the study findings (ns, no correlation; +,positive correlation).

(2008) incorporated two emotional expression conditions (positive and negative) into their design. They reported significant correlations between static and dynamic presentations of female images, similar to Lander (2008), and these occurred irrespective of emotional expression. However, they found positive but non-significant correlations between presentations of male images, in either emotional condition. Finally, Roberts et al. (2009) found a significant and positive correlation using female judgements of male images.

In view of these mixed results, we aimed to provide a further test of this interesting and important issue. We reasoned that the extent to which perceptions either do or do not differ under static/dynamic presentation may depend on methodological issues that require further exploration (summarised in Table 1). For example, both Rubenstein and Lander used a between-subjects design: rather than comparing static/dynamic judgements of female faces by the same raters, different sets of raters judged either condition (to some extent the same was true in Penton-Voak & Chang's study, where raters saw all faces but only one quarter of them in each of the four conditions). While using different raters is a reasonable step to ensure independence of ratings under two different conditions, well-documented individual differences in preferences (e.g. Little & Perrett 2002) could have, at least partially, obscured the relationship between the two measures. Furthermore, in Rubenstein's study, ratings were averaged across raters of both sexes; hence, ratings could have been confounded by the same-sex judgements made by females compared with the opposite-sex judgements made by males. Third, the context under which dynamic stimuli are captured might be influential for rater perceptions. Both Rubenstein and Lander video-recorded targets while reading from a cue card, while those in Penton-Voak & Chang's study were filmed both under a neutral context similar to using a cue card (reciting a series of numbers or pictures) and also while describing plans for a holiday where the stimuli were more naturalistic. In this vein, Roberts et al. (2009) asked men to film themselves in a scenario in which they were were asked to introduce themselves as although they were meeting an attractive woman. Finally, Penton-Voak & Chang (2008) and Roberts et al. (2009) presented videos without sound, controlling for variability in semantic content but also withholding from raters any vocal cues of mate quality which are known to influence attractiveness judgements (Feinberg et al. 2005a,b; Saxton et al. 2006). The same was true of Lander's (2008) study (pers.comm. to TKS) but it is unclear from Rubenstein (2005) whether or not dynamic stimuli were played with sound; if they were, vocal cues would have been available and might have influenced the judgements of dynamic cues in ways not possible under static presentation.

Here, we examined these issues by collecting image sets and ratings to compare static–dynamic correlations when judgements are made by the same or different sets of raters, by raters of the same- or opposite-sex to the targets, and when (silent) dynamic stimuli are collected under different scenarios.

Methods

Stimuli

The study was approved by the University of Liverpool's Committee on Research Ethics. Twenty male (mean age = 26.1) and 20 female (mean age = 22.2) undergraduates were photographed (Canon Powershot) standing, with a neutral expression, looking straight at the camera, in front of a plain background and in a windowless room lit with standard fluorescent lighting. Photographs were cropped just above the top of the head and to just below the waist, and normalised for horizontal height (Psychomorph; Tiddeman et al. 2001) to standardise presented image size. Images were resampled to 400×480 pixels (resolution 72 dpi). Seated participants were then video-recorded while they introduced themselves as they might to someone in a bar ('mate choice context'). Video clips were subsequently processed (Adobe After Effects 7.0) and edited to a duration of 20 s, cropped to dimensions of 400×480 square pixels and encoded as 25fps QuickTime movies using the MPEG-4 codec. For comparison of the context under which videos were recorded, an independent set of 20 males also participated; these were treated in exactly the same way as above, but during filming they described their most recent holiday (non-mate choice context). In both video contexts, videos were presented without sound to withhold semantic content and vocal cues from the raters, thus ensuring that any potential differences in judgements were solely because of visual cues.

Image Rating

For the main study, in which videos were recorded for the mate choice context, photographs and videos were rated by 96 raters in total. These comprised

two independent sets of 24 males and 24 females; one set saw photos first (mean age: males = 24.4, females = 22.0), the other saw videos first (mean age: males = 22.5, females = 22.1). Ratings were carried out using a 7-point scale anchored by the descriptors 'very unattractive' and 'very attractive'. In the non-mate choice context, another independent set of 10 females provided ratings.

Analyses

Mean ratings per target were analysed using Pearson's correlations. First, between-subjects correlations of static and dynamic image ratings were calculated for male and female stimuli separately, using either all raters (male and female combined) or only male and female raters. To obtain these between-subjects ratings in a way that controlled for exposure to the images, we used static image ratings from those raters who saw static images first and dynamic image ratings from raters who saw dynamic images first. Corresponding estimates of variance (r^2) in dynamic ratings that were explained by static ratings were computed through bivariate linear regression.

Second, we compared several experimental design effects on the strength of static–dynamic correlations using a simple resampling procedure to generate sets of correlation coefficients based on equally sized subsets of raters. We were able to compare the effect of within-subjects vs. between-subjects ratings because each rater saw images in each of the two presentation formats. Between-subjects ratings were obtained by comparing the ratings of static images by half the raters with the ratings of dynamic images by the other half. However, any possible subdivision of raters will likely generate a slightly different correlation coefficient, and, if it was done only once, it is theoretically possible that a strongly positive correlation might have arisen by chance. The resampling procedure avoided this problem. We randomly subdivided raters so that half were used to calculate mean ratings for static images and the other half were used to calculate mean ratings for dynamic images, then calculated the resulting static–dynamic correlation coefficient, and repeated this procedure to a total of 40 iterations. We also used the same approach for within-subjects ratings, so that the power of the correlations was equivalent across all relevant comparisons. We calculated 40 iterations for each of eight comparisons (image: male/female; rating design: within/between subjects; rater sex: opposite/same). The resulting 320 correlation coefficients were then analysed by factorial ANOVA (data fulfilled assumptions of this test). This analysis was performed separately for raters who saw photos first and those who saw videos first.

Finally, we compared the effect of context (mate choice vs. non-mate choice task during collection of dynamic images) by comparing two corresponding sets of 40 iterated correlation coefficients for male static and dynamic images, based on independent sets of raters and stimuli (i.e. the within-subjects design with equivalent numbers of raters), using an independent-sample t-test.

Results

Correlations of mean ratings of static and dynamic images, for 20 male and 20 female targets, show that attractiveness judgements are strongly and positively correlated (Fig. 1, Table 2), regardless of whether they were based on judgements by raters of the same- or opposite-sex, or by all raters. In each case, correlations were significant with coefficients at least over 0.728 and with r^2 values of at least 0.531.

Using ANOVA with static/dynamic correlation coefficients as the dependent variable and stimulus sex, rater sex and rating design (within- or between-subjects) as factors, we found significant main effects of rating design, rater sex and stimulus sex, and some interactions (Table 3). Regardless of whether static or dynamic images were seen first, higher correlations were found using a within-subject rating design. However, other effects indicated that the strength and direction of differences in correlations varied according to the order of image presentation. When static images were seen first, correlations between ratings of static and dynamic images were higher for judgements of male than female images, and for same-sex than opposite-sex ratings (see Fig. 2). The opposite was true when videos were seen first. Additionally, there were significant interactions between rating design and rater sex (correlations were lowest in opposite-sex ratings in the between-subjects design, indicating this was where lies the lowest degree of concordance in ratings), and between stimulus sex and rater sex: correlations between male static and dynamic images were lower when judged by females than males (at least in the photo-first order), and correlations between female static and dynamic images were higher when judged by males than females (at least in the video-first order).

Finally, we investigated the effects of context, using an independent set of raters and stimuli.
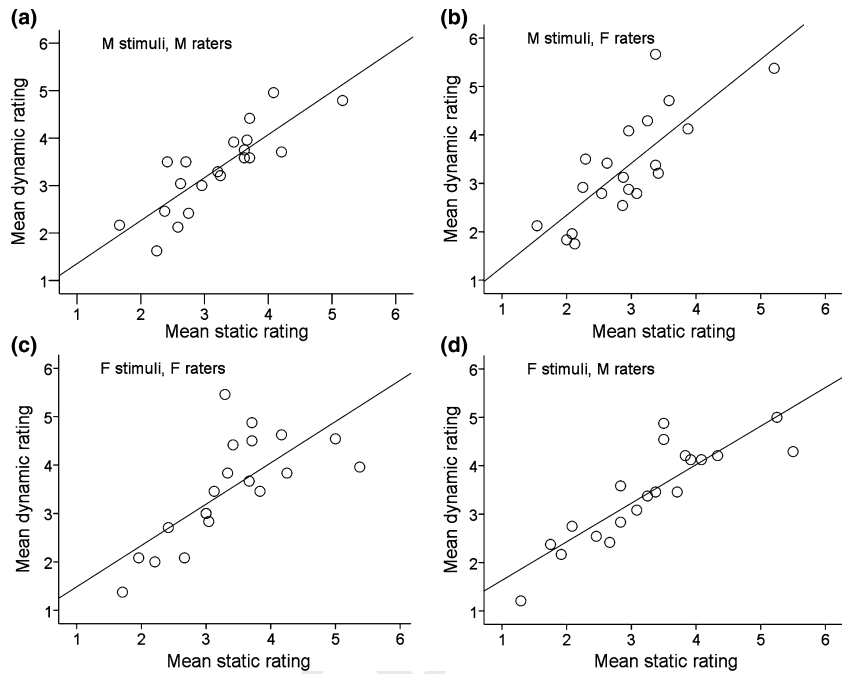
Fig. 1: Relationship between mean ratings of static and dynamic facial images of 20 males (a, b) and 20 females (c, d). In each panel, mean static ratings and mean dynamic ratings are calculated from independent sets of 24 raters who had not yet seen the images in the other format. In (a, c), raters are of the same sex as the stimuli, in (b, d) they are of the opposite sex to the stimuli. For more details, see text and Table 2.

Table 2: Relationships between male and female static and dynamic stimuli when judged by same-sex or opposite-sex raters, or both male and female raters together

Stimuli	Same-sex raters	Opposite-sex raters	All raters
Male	0.838 (0.702), <0.001	0.785 (0.617), <0.001	0.834 (0.696), <0.001
Female	0.728 (0.531), <0.001	0.866 (0.750), <0.001	0.809 (0.654), <0.001

Values are given as Pearson's r (r^2) and p.

Table 3: Results of ANOVA on Pearson's correlation coefficients

Effect	F	df	p
Photo-first order			
Stimulus sex	83.24	1, 312	<0.001
Rating design (WS vs. BS)	19.83	1, 312	<0.001
Rater sex	146.38	1, 312	<0.001
Stimulus sex × rater sex	112.14	1, 312	<0.001
Rating design × rater sex	7.51	1, 312	0.006
Video-first order			
Stimulus sex	42.96	1, 312	<0.001
Rating design (WS vs. BS)	48.56	1, 312	<0.001
Rater sex	58.61	1, 312	<0.001
Stimulus sex × rater sex	43.62	1, 312	<0.001
Rating design × rater sex	0.075	1, 312	0.784

Rating design denotes whether static and dynamic images were judged by same (WS, within-subjects) or different (BS, between-subjects) rater sets. Other interaction effects were not statistically significant.

Within-subject ratings for static and dynamic images in the non-mate choice context (description of recent holiday) were again strongly correlated (Pearson's $r = 0.836$, p < 0.001). However, comparison of female ratings for males video-recorded in the mate choice and non-mate choice contexts revealed that the distribution of iterated correlation coefficients between contexts were significantly different, being higher in the mate choice context ($\bar{x} \pm$ SE; non-mate choice = 0.704 ± 0.011, mate choice = 0.741 ± 0.013; $t_{78} = 2.13$, $p = 0.036$).

Discussion

Our results indicate that ratings of static and dynamic images can be strongly positively correlated. Importantly, the strength of this correlation is dependent on (1) whether the two sets of ratings are produced by the same or different raters, (2) whether raters are of the same- or opposite-sex to targets and (3) the context in which dynamic images are collected.

Previous studies have found mixed evidence for the strength of this relationship, although in all cases the relevant correlations are positive, even if only weakly so (Rubenstein 2005; Lander 2008; Penton-Voak & Chang 2008). To summarise the specific
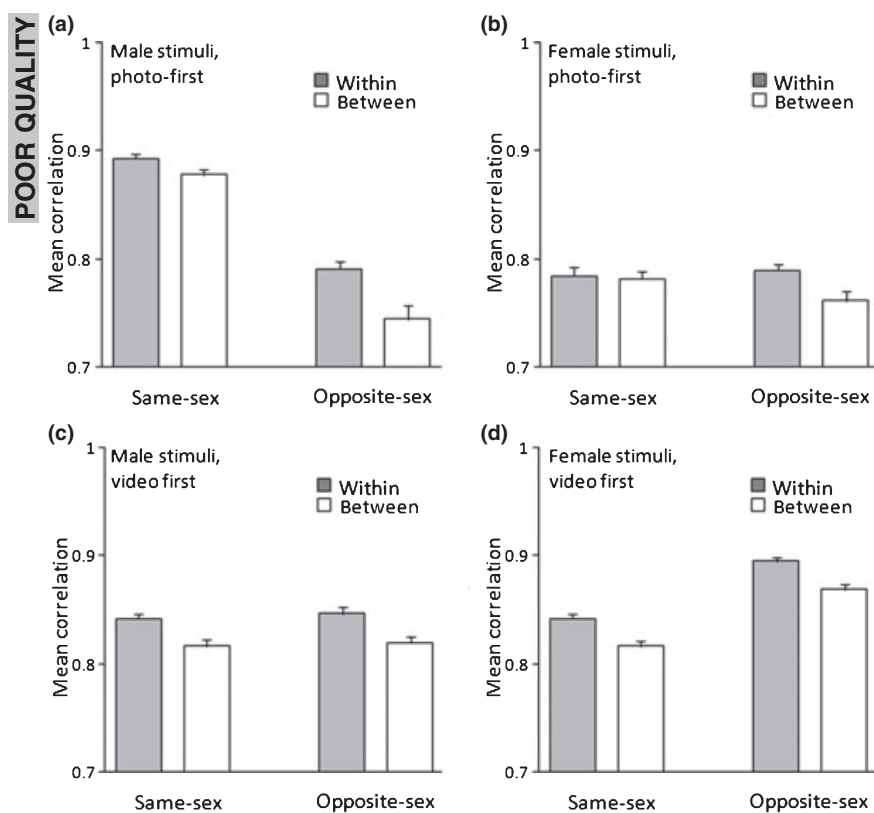
Fig. 2: Mean (± SE) Pearson's correlation coefficients from static–dynamic image correlations. Each bar represents means from 40 iterations of randomised subsamples of raters. Within-subject means reflect correlation strength using the same raters (same-sex or opposite-sex to targets); between-subject means reflect correlations between static ratings by one rater set and dynamic ratings by another rater set. See Table 3 for significance levels of main effects and interactions.

findings from these studies (see also Table 1), significant static–dynamic correlations have been found for female faces by Lander and Penton-Voak & Chang but not by Rubenstein. For male images, no significant correlations were found by either Lander or Penton-Voak & Chang (Rubenstein did not use male images), although Roberts et al. did report a significant positive relationship. Our subsequent analyses revealed that the lowest correlations were found in female ratings of males (although only when photos were presented first, Fig. 2 and Table 3), which is somewhat consistent with both Lander (2008) and Penton-Voak & Chang (2008). These results indicate that there may be sex-specific differences in the extent to which static images reflect judgements of mate value in more ecologically valid contexts (see also Penton-Voak & Chang 2008). This may be because attributions of personality appear more salient in opposite-sex judgements of males than of females (Berry & Miller 2001), a possibility supported by our finding that the correlation was least strong amongst female raters, and weaker than opposite-sex judgements made by male raters. However, this same effect was not evident when videos were rated before photos, suggesting that, having seen videos, women were influenced by remem-

bered impressions based on dynamic cues when rating photos (see also Lander & Bruce 2000).

Each of the methodological refinements we included in this study might at least partly account for the discrepancies in results between studies and the more positive relationships that we report here. Adoption of a within-subjects rating design is one important aspect that should be easy to incorporate into future studies and it is a logical one if the researcher's interest is to understand how individual ratings differ across two conditions. However, we note that one weakness of this is that there may be carry-over effects from judgements made under one condition and those under the other. Indeed, although we found robust positive correlations in all comparisons, we did find slightly different correlation strengths depending on order of image presentation (notably with respect to female ratings of male images). If the same raters were to be used, then presenting static images before dynamic ones is arguably more consistent with the natural progression that occurs during a first-time encounter with an unfamiliar individual, where an initial impression is made within as little as 100 ms (Willis & Todorov 2006) and this is presumably more weighted by structural facial features contained within static

images; behavioural cues and mannerisms such as those evident in a 20 s video are likely to manifest after this first impression.

Furthermore, our results showing rater sex and stimulus sex main effects and interactions suggest that studies examining rated attractiveness should, wherever possible, choose either male or female raters, as appropriate to the needs of the study, rather than conflate male and female raters in one assessment set. Finally, the context under which the dynamic images are recorded also affected the strength of the correlations reported and so should be carefully chosen when planning experiments. Here, we found higher static/dynamic correlations when participants were asked to introduce themselves than we did when they were asked to speak about a recent holiday. This may be because these different contexts elicit different or more variable sets of expressions and movement cues. It is possible that self-introduction is perceived as a task that is more difficult than recounting a series of previous events, and that this difficulty elicits behavioural cues which more reliably reflect mate quality than those evinced by other contexts. This interpretation is also consistent with the low correlations recorded during an emotionally neutral and relatively easier task such as reading from a cue card (Rubenstein 2005; Lander 2008). We conclude that context is likely to be an important aspect of design, but one which requires more investigation.

We certainly do not dispute that dynamic images contain additional information for facial perception to that contained in static images, including salient emotional cues. We also agree that dynamic cues may be particularly informative for face recognition (e.g. Lander & Chuang 2005) or perception of emotion, personality and attractiveness (e.g. Rubenstein 2005; Morrison et al. 2007; Roberts 2008), notwithstanding the contributions also made by features contained within neutral-posed, static images (e.g. Penton-Voak et al. 2006; Little & Perrett 2007; Roberts & Little 2008). However, in this study, most (53–75%) of the variance in dynamic ratings is explained by static ratings, even when using a between-subjects design. An outstanding problem for future research using dynamic images, therefore, is how to investigate the contribution of dynamic cues independently of information contained in static appearance. We believe that solutions would include using residuals from regression with dynamic ratings as the dependent variable (such residuals would represent the unique contribution of movement to face perception: see for example,

Roberts et al. 2009) or using static image ratings as a covariate to control for their effect. We suggest that these approaches could prove useful in attempts to disentangle and evaluate the role of dynamic information in face perception.

In conclusion, although a fuller explanation of the role of dynamic information in human attractiveness judgements remains an intriguing area of enquiry, our data suggest that the disparity between judgements in image presentation type may be far smaller than previously suggested: the attractiveness of a static face is a good predictor of that person's attractiveness when moving. In this way, static and dynamic stimuli cue similar attractiveness judgements.

Acknowledgements

Literature Cited

Berry, D. S. & Miller, K. M. 2001: When boy meets girl: attractiveness and the five-factor model in opposite-sex interactions. J.Res.Pers. **35**, 62—77.

Conway, C. A., Jones, B. C., Debruine, L. M. & Little, A. C. 2008: Evidence for adaptive design in human gaze preference. Proc. Roy. Soc. B **275**, 63—69.

DeBruine, L. M., Jones, B. C. & Perrett, D. I. 2005: Women's attractiveness judgments of self-resembling faces change across the menstrual cycle. Horm. Behav. **47**, 379—383.

Feinberg, D., Jones, B., Little, A., Burt, D. & Perrett, D. 2005a: Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. Anim. Behav. **69**, 561—568.

Feinberg, D. R., Jones, B. C., DeBruine, L. M., Moore, F. R., Law Smith, M. J., Cornwell, R. E., Tiddeman, B. P., Boothroyd, L. G. & Perrett, D. I. 2005b: The voice and face of woman: one ornament that signals quality? Evol. Hum. Behav. **26**, 398—408.

Gangestad, S. W. & Scheyd, G. J. 2005: The evolution of human physical attractiveness. Ann. Rev. Anthropol. **34**, 523—548.

Gangestad, S. W., Simpson, J. A., Cousins, A. J., Garver-Apgar, C. E. & Christensen, P. N. 2004: Women's preferences for male behavioral displays change across the menstrual cycle. Psychol. Sci. **15**, 203—207.

Grammer, K. 1990: Strangers meet – laughter and non-verbal signs of interest in opposite-sex encounters. J. Nonverb. Behav. **14**, 209—236.

Grammer, K., Kruck, K., Juette, A. & Fink, B. 2000: Non-verbal behavior as courtship signals: the role of control and choice in selecting partners. Evol. Hum. Behav. **21**, 371—390.

Grammer, K., Fink, B., Moller, A. P. & Thornhill, R. 2003: Darwinian aesthetics: sexual selection and the biology of beauty. Biol. Rev. **78**, 385—407.

Jones, B. C., Little, A. C., Boothroyd, L., Feinberg, D. R., Cornwell, R. E., DeBruine, L. M., Roberts, S. C., Penton-Voak, I. S., Smith, M. J. L., Moore, F. R., Davis, H. P. & Perrett, D. I. 2005: Women's physical and psychological condition independently predict their preference for apparent health in faces. Evol. Hum. Behav. **26**, 451—457.

Lander, K. 2008: Relating visual and vocal attractiveness for moving and static faces. Anim. Behav. **75**, 817—822.

Lander, K. & Bruce, V. 2000: Recognizing famous faces: exploring the benefits of facial motion. Ecol. Psychol. **12**, 259—272.

Lander, K. & Chuang, L. 2005: Why are moving faces easier to recognize? Vis. Cogn. **12**, 429—442.

Little, A. C. & Perrett, D. I. 2002: Putting beauty back in the eye of the beholder: evolution and individual differences in face preference. Psychologist **15**, 28—32.

Little, A. C. & Perrett, D. I. 2007: Using composite face images to assess accuracy in personality attribution. Br. J. Psychol. **98**, 111—126.

Little, A. C., Burt, D. M., Penton-Voak, I. S. & Perrett, D. I. 2001: Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. Proc. Roy. Soc. B **268**, 39—44.

Morrison, E. R., Gralewski, L., Campbell, N. & Penton-Voak, I. S. 2007: Facial movement varies by sex and is related to attractiveness. Evol. Hum. Behav. **28**, 186—192.

Penton-Voak, I. S. & Chang, H. Y. 2008: Attractiveness judgements of individuals vary across emotional expression and movement conditions. J. Evol. Psychol. **6**, 89—100.

Penton-Voak, I. S. & Perrett, D. I. 2000: Female preference for male faces changes cyclically: further evidence. Evol. Hum. Behav. **21**, 39—48.

Penton-Voak, I. S., Perrett, D. I. & Pierce, J. 1999: Computer graphic studies of the role of facial similarity in attractiveness judgements. Curr. Psychol. **18**, 104—117.

Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B. P., Burt, D. M. & Perrett, D. I. 2001: Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. Proc. Roy. Soc. B **268**, 1617—1623.

Penton-Voak, I. S., Pound, N., Little, A. C. & Perrett, D. I. 2006: Personality judgments from natural and composite facial images: more evidence for a ''kernel of truth'' in social perception. Soc. Cogn. **24**, 490—524.

Perrett, D. I., May, K. A. & Yoshikawa, S. 1994: Facial shape and judgments of female attractiveness. Nature **368**, 239—242.

Perrett, D. I., Lee, K. J., Penton-Voak, I. S., Rowland, D. R., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D. L. & Akamatsu, S. 1998: Effects of sexual dimorphism on facial attractiveness. Nature **394**, 884—887.

Rhodes, G. 2006: The evolutionary psychology of facial beauty. Ann. Rev. Psychol. **57**, 199—226.

Roberts, S. C. 2008: Nonverbal communication of mate quality in humans. In: Sociobiology of Communication: An Interdisciplinary Perspective (D'Ettorre, P. & Hughes, D. P., eds). Oxford Univ. Press, Oxford, pp. —.

Roberts, S. C. & Little, A. C. 2008: Good genes, complementary genes and human mate choice. Genetica **132**, 309—321.

Roberts, S. C., Havlicek, J., Flegr, J., Hruskova, M., Little, A. C., Jones, B. C., Perrett, D. I. & Petrie, M. 2004: Female facial attractiveness increases during the fertile phase of the menstrual cycle. Proc. Roy. Soc. B (Suppl.) **271**, S270—S272.

Roberts, S. C., Little, A. C., Gosling, L. M., Perrett, D. I., Carter, V., Jones, B. C., Penton-Voak, I. & Petrie, M. 2005a: MHC-heterozygosity and human facial attractiveness. Evol. Hum. Behav. **26**, 213—226.

Roberts, S. C., Little, A. C., Gosling, L. M., Jones, B. C., Perrett, D. I., Carter, V. & Petrie, M. 2005b: MHC-assortative facial preferences in humans. Biol. Lett. **1**, 400—403.

Roberts, S. C., Little, A. C., Lyndon, A., Roberts, J., Havlicek, J. & Wright, R. L. 2009: Manipulation of body odour alters men's self-confidence and judgements of their visual attractiveness by women. Int. J. Cosmet. Sci. **31**, 47—54.

Rubenstein, A. J. 2005: Variation in perceived attractiveness – differences between dynamic and static faces. Psychol. Sci. **16**, 759—762.

Saxton, T. K., Caryl, P. G. & Roberts, S. C. 2006: Vocal and facial attractiveness judgments of children, adolescents and adults: the ontogeny of mate choice. Ethology **112**, 1179—1185.

Thornhill, R. & Gangestad, S. W. 1999: Facial attractiveness. Trends Cogn. Sci. **3**, 452—460.

Tiddeman, B., Burt, M. & Perrett, D. 2001: Prototyping and transforming facial textures for perception research. IEEE Comput. Graph. Appl. **21**, 42—50.

Willis, J. & Todorov, A. 2006: First impressions: making up your mind after a 100-ms exposure to a face. Psychol. Sci. **17**, 592—598.

Author Query Form

Dear Author,
During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

Query reference	Query	Remarks
Q1	**AUTHOR: Please supply a short title that will be used as the running head.**	
Q2	**AUTHOR: Rubenstein (2006) has been changed to Rubenstein (2005) so that this citation matches the Reference List. Please confirm that this is correct.**	
Q3	**AUTHOR: Please provide the page range for reference Roberts (2008).**	
Q4	**AUTHOR: Figure 2 is of poor quality. Please check required artwork specifications at http://www.blackwellpublishing.com/authors/digill.asp**	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	Textual mark	Marginal mark
Leave unchanged	· · · under matter to remain	⊘
Insert in text the matter indicated in the margin	⋏	New matter followed by ⋏ or ⋏⊗
Delete	/ through single character, rule or underline or ├──────┤ through all characters to be deleted	⌿ or ⌿⊗
Substitute character or substitute part of one or more word(s)	/ through letter or ├──────┤ through characters	new character / or new characters /
Change to italics	— under matter to be changed	⌣
Change to capitals	≡ under matter to be changed	≡
Change to small capitals	= under matter to be changed	=
Change to bold type	∿ under matter to be changed	∿
Change to bold italic	≁ under matter to be changed	≁
Change to lower case	Encircle matter to be changed	≢
Change italic to upright type	(As above)	⥮
Change bold to non-bold type	(As above)	⥮
Insert 'superior' character	/ through character or ⋏ where required	Y or ⋋ under character e.g. Y² or ⋋²
Insert 'inferior' character	(As above)	⋏ over character e.g. ⋏₂
Insert full stop	(As above)	⊙
Insert comma	(As above)	,
Insert single quotation marks	(As above)	Y or ⋋ and/or Y or ⋋
Insert double quotation marks	(As above)	Y or ⋋ and/or Y or ⋋
Insert hyphen	(As above)	⊢⊣
Start new paragraph	⌐	⌐
No new paragraph	↩	↩
Transpose	⊔⊓	⊔⊓
Close up	linking ◡ characters	◯
Insert or substitute space between characters or words	/ through character or ⋏ where required	Y
Reduce space between characters or words	│ between characters or words affected	↑