

This is a repository copy of *Facial and self-report questionnaire measures capture different aspects of romantic partner preferences*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/136757/>

Version: Accepted Version

Article:

South Palomares, Jennifer Kay and Young, Andrew William orcid.org/0000-0002-1202-6297 (2018) Facial and self-report questionnaire measures capture different aspects of romantic partner preferences. *British journal of psychology*. ISSN 0007-1269

<https://doi.org/10.1111/bjop.12347>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Authors' version, as accepted for publication in *British Journal of Psychology*. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article.

Date of acceptance: 16th August 2018.

Facial and self-report questionnaire measures capture different aspects of romantic partner preferences

Jennifer K. South Palomares and Andrew W. Young

Department of Psychology, University of York, Heslington, York YO10 5DD, UK

Acknowledgements: Data reported here formed part of JSP's PhD thesis (University of York).

Correspondence should be addressed to:

Jennifer K. South Palomares, Department of Education, University of York, Heslington, York YO10 5DD, UK, jennifer.south.palomares@york.ac.uk

Keywords

Face perception; First impressions; Person perception; Partner preferences; Romantic relationships.

Abstract

Romantic relationship researchers often use self-report measures of partner preferences based on verbal questionnaires. These questionnaires show that partner preferences involve an evaluation in terms of underlying factors of vitality-attractiveness, status-resources, and warmth-trustworthiness. However, when people first encounter a potential partner, they can usually derive a wealth of impressions from their face, and little is known about the relationship between verbal self-reports and impressions derived from faces. We conducted four studies investigating potential parallels and differences between facial impressions and verbal self-reports. Study 1 showed that when evaluating highly variable everyday face images in a context that does not require considering them as potential partners, participants can reliably perceive the traits and factors that are related to partner preferences. However, despite being capable of these nuanced evaluations, Study 2 found that when asked to evaluate images of faces as potential romantic partners, participants' preferences become dominated by attractiveness-related concerns. Study 3 confirmed this dominance of facial attractiveness using morphed face-like images. Study 4 showed that attractiveness dominates partner preferences for faces even when task instructions imply that warmth-trustworthiness or status-resources should be of primary importance. In contrast to verbal questionnaire measures of partner preferences, then, evaluations of faces focus heavily on attractiveness, whereas questionnaire self-reports tend on average to prioritise warmth-trustworthiness over attractiveness. Evaluations of faces and verbal self-report measures therefore capture different aspects of partner preferences.

Introduction

From self-report questionnaires, romantic relationship researchers have identified a tripartite structure underlying verbally expressed partner preferences in which desired traits correspond to factors involving vitality-attractiveness, status-resources, and warmth-trustworthiness (Fletcher, Simpson, Thomas, & Giles, 1999). These factors can to some extent be flexibly used, in the sense that one individual might prioritise finding a partner who is warm and trustworthy while another individual will prioritise vitality and attractiveness and someone else seeks status and resources. On average, however, people tend consistently to value warmth-trustworthiness (Buss, 1989; Fletcher et al., 1999; Fletcher, Tither, O'Loughlin, Friesen, & Overall, 2004; Li, Bailey, Kenrick, & Linsenmeier, 2002). The underlying three-factor structure has been widely validated (Campbell, Simpson, Kashy, & Fletcher, 2001; Fletcher, Kerr, Li, & Valentine, 2014; Fletcher, Simpson, & Thomas, 2000; Fletcher et al., 1999; Fletcher & Simpson, 2000) and shown to have predictive validity for relationship outcomes (Campbell, Chin, & Stanton, 2016; Campbell et al., 2001; Fletcher et al., 2000, 1999).

Although relationship researchers often examine verbally expressed and self-reported partner preferences (e.g., Buss & Barnes, 1986; Fletcher et al., 1999; Simpson & Gangestad, 1992), many relationships begin in contexts where facial impressions offer an important potential source of information. For instance, many individuals maintain an online profile for romance (e.g., the online dating website eHarmony boasts almost 2.5 million UK members; eHarmony, 2015), which typically involves posting facial photographs. Indeed, the prevalence of online images and internet-based relationships make the questions we address particularly timely and important.

Facial evaluations are known to influence such significant events as political elections, online financial lending, and court decisions (Olivola, Funk, & Todorov, 2014).

They have also been shown to affect romantic preferences (Fiore, Taylor, Mendelsohn, & Hearst, 2008; Fletcher et al., 2014; Hancock & Toma, 2009; Todd, Penke, Fasolo, & Lenton, 2007), but relatively little is understood concerning how such influences operate and how they might relate to preferences assessed from self-report questionnaires. We reasoned that facial impressions may be particularly consequential in online contexts and on first acquaintance, given the sparse additional information typically available to viewers. Importantly, we sought to establish whether partner preference evaluations derived from facial and verbal (questionnaire) sources of information are consistent with each other.

In particular, the partner preference model based on verbal self-report (Fletcher et al., 1999) might not apply to facial judgements if verbal (i.e. conceptual) and facial (perceptual) evaluations are derived in different ways. Examining whether a leading partner preference model based on self-report questionnaire items (Fletcher et al., 1999) can be applied to first impressions of faces is relevant to understanding its potential applicability to partner selection in everyday situations which include facial information.

Because rapid evaluations of faces are common in online contexts, we grounded our approach in the facial first impressions literature (Sutherland et al., 2013; Sutherland, Oldmeadow, & Young, 2016; Todorov, 2017; Vernon, Sutherland, Young, & Hartley, 2014). In previous studies, research has found that traits from a leading model of self-reported partner preferences (Fletcher et al., 1999) are readily evaluated in naturally occurring, highly variable, face images of the type often encountered in daily life (South Palomares, Sutherland, & Young, 2018; South Palomares & Young, 2017). Moreover, research has been able to demonstrate that when evaluated from face images these traits reflected an underlying factor structure that closely resembled both Fletcher et al.'s (1999) verbal self-report partner preferences model and a more general model of facial first impressions (South Palomares et al., 2018; Sutherland et al., 2013, 2016; Vernon et al., 2014).

Although these previous studies by South Palomares et al. (2018) and South Palomares and Young (2017) show that people are, in principle, able to make evaluations of faces that could correspond to their verbally expressed partner preferences, they did not directly test this potential correspondence. Instead, participants were asked to evaluate face images in contexts where partner preferences as such were not explicitly requested. In the present research, we began in the same way (Study 1) by creating a factor structure corresponding to how 12 traits derived directly from Fletcher et al.'s (1999) model of verbal self-report questionnaire items were seen in 1,000 highly varied face images. We then used the data from Study 1 to investigate the relationship between facial and verbal measures of partner preferences (in Studies 2 to 4) by examining how explicit judgements of partner preferences from face images related to (1) the trait structure found in Study 1 and (2) verbally expressed preferences from Fletcher et al.'s (1999) verbal self-report questionnaire. Facial attractiveness was found to dominate the evaluations made by nearly all participants across Studies 2 to 4, showing a clear difference between measures of partner preference derived from facial or verbal self-report sources of information.

Study 1

Study 1 involved participants' ratings of 1,000 everyday face images on one of 12 traits selected to represent Fletcher et al.'s (1999) verbal self-report partner preference factors. Ratings of 10 of these 12 traits came from previous work (South Palomares & Young, 2017) and on this basis we expected to be able to capture an underlying three-factor structure. However, to ensure that Study 1 offered the closest possible comparison to Fletcher et al.'s (1999) model we selected the 12 traits for Study 1 by taking the two traits corresponding to each of the three Fletcher et al. factor labels (e.g., warmth and trustworthiness for the warmth-trustworthiness factor) and the next two traits loading most highly on their respective verbal factors, leading to a list of four traits for each of the three

verbal factors (sexy, attractiveness, vitality, and adventurous representing vitality-attractiveness; has a good job, has status, has resources, and is financially secure representing status-resources; trustworthiness, supportive, understanding, and warmth representing warmth-trustworthiness).

Many studies in face perception begin by standardising the stimuli (e.g. for lighting, pose, expression) with the aim of eliminating image differences considered to form unwanted ‘noise’. Our approach instead aimed to capture responses to highly variable faces of the type we encounter in our everyday lives (termed ambient images by Jenkins, White, Van Montfort, & Burton, 2011). Ambient face images represent a novel approach to examining first impressions of potential partners. Because ambient images maintain the highly variable properties that are present in natural environments, they can allow findings to be more directly generalised to real life contexts, relative to controlled stimuli (Sutherland, Rhodes, & Young, 2017; Young, 2018).

In line with previous studies seeking to model the underlying factor structure of facial first impressions (South Palomares et al., 2018; Sutherland et al., 2013, 2016; Vernon et al., 2014) we used the averaged responses to each of the 1,000 face images as the primary data for Study 1, rather than individual participants’ responses. Recent work has shown that individual participant responses comprise a mix of ‘shared’ (common to all participants) and ‘private’ (idiosyncratic to that participant) taste (Germine et al., 2015; Hehman, Sutherland, Flake, & Slepian, 2017; Hönekopp, 2006). By averaging participants’ responses, we reduce much of the impact of differences in private taste and can thus establish whether there is an underlying factor structure based on shared taste. Study 1 formed an essential starting point that was used to create a set of averaged trait ratings for each face image and a set of weightings that could correspond to each of the Fletcher et al. (1999) factors. In Studies 2 to 4, which are concerned with the applicability of this underlying structure at the individual

participant level, we return to the more conventional technique of analysing each participant's responses.

The data for Studies 1 to 4 are available as supplementary material.

Method

Participants. The study involved data from separate groups of 10 participants (five male, five female) rating 1,000 face images on one of 12 traits (see below), in order to arrive at an average rating (across the 10 raters) for each of the 12 traits for each image. Only a single trait was evaluated by each participant to avoid carry-over effects in which correlations between trait judgements can be inflated by participants being influenced by their previous judgements (Rhodes, 2006). Ratings of 10 of the 12 traits involved were already available from previous studies (South Palomares et al., 2018; South Palomares & Young, 2017; Sutherland et al., 2013, 2016). Twenty further participants, university students, were recruited via the University (50% male, mean age of 22 years, $SD = 3.45$) to create the groups of 10 participants needed to rate each of the remaining two traits ('vitality' and 'has resources'). Participants were self-reported native English speakers, raised in a Western environment. The participants provided written consent to procedures approved by the Ethics Committee of the University Psychology Department. Participants did not take part in the other reported studies. **Face images.** The study used a database of 1,000 face images (Santos & Young, 2005, 2008, 2011), representing 500 female and 500 male Caucasian non-famous adults. Like other face databases (e.g., Oosterhof & Todorov, 2008), to avoid other-race effects (O'Toole & Natu, 2013), non-Caucasian faces were not included. The images were taken from the Internet and were deliberately unconstrained in terms of their variability (e.g., differences in pose, lighting, background, age, expression, and facial hair, amongst others). All variables excepting Caucasian adult appearance were unstandardised. Images were cropped to reveal only the individuals' head and shoulders and standardised to 150 pixels in height. See Figure

1 for examples of the type of images used.



Figure 1. Example ambient images like those used in the study, from the authors' personal collections. Photographs from the database are not shown for copyright reasons.

Partner preference trait ratings. Participants rated images on one of 12 traits representing the warmth-trustworthiness, status-resources, and vitality-attractiveness factors from Fletcher et al.'s (1999) partner preference model. Four traits were chosen to represent each factor; two of these were the factor labels (e.g., warmth and trustworthiness for the warmth-trustworthiness factor) and a further two traits loaded highest on their respective verbal factors. The resulting 12 traits were as follows: sexy, attractiveness, vitality, and adventurous (representing vitality-attractiveness); has a good job, has status, has resources, and is financially secure (representing status-resources, these were accompanied by '*or potential to achieve*' in parentheses as per Fletcher et al., 1999); and trustworthiness, supportive, understanding, and warmth (representing warmth-trustworthiness).

Procedure. Participants were informed that the study involved facial evaluations and that the task was self-paced, but to rely on their first impressions (Sutherland et al., 2013; Todorov, Mandisodza, Goren, & Hall, 2005). Participants completed six practice trials, rating faces randomly selected from the database, and then rated all 1,000 images, in a random order, on a single trait. Ten participants (five male) were randomly assigned to rate one of two traits: vitality and resources. Ratings for the remaining 10 traits were already available from previous studies using the same procedure (South Palomares et al., 2018; South Palomares & Young, 2017; Sutherland et al., 2013, 2016). Ratings were given for a specific trait (e.g, 'vitality'); partner preferences *per se* were not mentioned. Ratings were made on a

1-7 Likert scale. Images remained on the screen while participants made their judgement. The inter-trial interval was 750 ms. On completion, participants were debriefed and reimbursed with a small payment. The task was programmed using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, USA) and took 60 minutes.

Results

There was good internal consistency of trait judgements across raters for all 12 traits (Cronbach's alphas over .72; Nunnally & Bernstein, 1994), in line with previous studies (e.g. Sutherland et al., 2013). We do not seek to maintain that these evaluations are valid, given that previous studies (Jenkins et al., 2011; Sutherland, Young, & Rhodes, 2017; Todorov & Porter, 2014) have shown that evaluations are often image-specific (the same person can look trustworthy in one photograph and untrustworthy in another photograph). What matters here is that the internal consistency of ratings of each image allows meaningful factor analysis to probe the underlying structure.

We explored the underlying structure of the traits by creating an average rating for each trait from each image. Following the procedure used by Sutherland et al. (2013), a factor analysis was carried out on the ratings of the photographs to model the structure of the 12 partner preference traits as represented by face images. Factor analysis was preferred to principal components analysis (PCA) since it is better than PCA for model building as it attempts to model the structure between the variables and contains an estimation of error (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

The value for the Kaiser-Meyer-Olkin measure of sampling adequacy was .89 and Bartlett's test of Sphericity was $\chi^2(66) = 12875.26, p < .001$, indicating that a factor analysis was suitable. Regarding factor extraction, four criteria were used to be as objective as possible: Kaiser's criterion, the scree test (Fabrigar et al., 1999; Kline, 1994; O'Connor, 2000), a parallel analysis (Horn, 1965), and Velicer's (1976) minimum average partial

analysis. All four criteria revealed a three-factor solution. To examine the factor structure, a direct oblimin rotation on the principal axis factor analysis was conducted. An oblique rotation was used to avoid forcing the factors to be orthogonal and we interpreted loadings above .50. Table 1 shows the structure matrix, which depicts correlations between the factors and variables.

Table 1. Structure matrix of a principal axis factor analysis with direct oblimin rotation for trait ratings of faces in Study 1. The first four traits were derived from Fletcher et al.'s (1999) vitality-attractiveness factor, the next four represented status-resources, and the last four warmth-trustworthiness.

Trait	Factor 1	Factor 2	Factor 3
Sexy	.21	.31	.95
Attractiveness	.36	.33	.93
Vitality	.75	.39	.78
Adventurous	.63	.23	.71
Good job	.30	.81	.39
Status	.40	.90	.35
Resources	.35	.92	.30
Financially secure	.28	.92	.09
Trustworthiness	.88	.46	.42
Supportive	.92	.33	.36
Understanding	.87	.41	.30
Warmth	.93	.25	.19

Note: substantial loadings (over .50) are in bold.

Overall, factor one primarily represented what Fletcher et al. (1999) considered to be warmth-trustworthiness traits, factor two represented status-resources traits, and factor three represented vitality-attractiveness traits. The exceptions were that, whilst vitality and adventurous were traits selected to represent the vitality-attractiveness factor, they also loaded strongly on warmth-trustworthiness. The scale reliability for traits loading over .50 was calculated for each factor and revealed good reliability, with Cronbach's alphas over .90 for each factor (i.e., traits loading strongly on each factor did reliably represent these factors).

Following an oblique rotation, the variance accounted for by each factor cannot be

identified. However, to ascertain model robustness, we also conducted a PCA, which derived an almost identical three-factor structure and revealed that factor 1 accounted for 53% variance, factor 2 accounted for 19% variance, and factor 3 accounted for 15% variance. To further test model robustness, various other analyses were conducted, including a maximum likelihood factor analysis with promax rotation, an alpha factor analysis with direct oblimin rotation, a PCA with orthogonal rotation, and a split half analysis, which again derived almost identical three-factor structures. Hence, the three-factor solution is not dependent on the specific analysis conducted. Furthermore, since the traits vitality and adventurous loaded strongly on two factors in the three-factor solution, a principal axis factor analysis with a direct oblimin rotation to extract four factors was conducted to determine whether a four-factor solution would be more adequate. The fourth factor accounted for 1% of the variance and contained only one trait loading over .50. Therefore, it seems reasonable to infer that a three-factor solution best represents the traits selected to correspond with Fletcher et al.'s (1999) verbal questionnaire partner preference factors.

Study 1 used traits intended directly to represent Fletcher et al.'s (1999) partner preference model, by selecting traits representing the factor labels of this model (i.e., warmth and trustworthiness for the warmth-trustworthiness factor) alongside other high loading traits. A previous study by South Palomares et al. (2018) already created a model using traits taken from Fletcher et al.'s (1999) model, but without including the factor labels as traits. Hence, it is relevant to examine the overlap between the structure of facial impressions of traits from Study 1 and the face-related factor weightings from South Palomares et al. (2018). To make this comparison, a factor score for each of the 1,000 images was calculated to represent its loadings on each of the three factors of the face-related partner preference model derived from the Study 1 factor analysis. These image-specific factor scores were correlated with equivalent factor scores indexing the three face-related factors for the same 1,000

photographs taken from South Palomares et al. (2018). High correlations ($r_s > .96$, $p_s < .001$, $n_s = 1,000$) were found between (1) the warmth-trustworthiness factors, (2) the status-resources factors, and (3) the vitality-attractiveness factors (see Figure 2). These data are sufficient to establish a strong correspondence between the factors underlying these two face-related models.

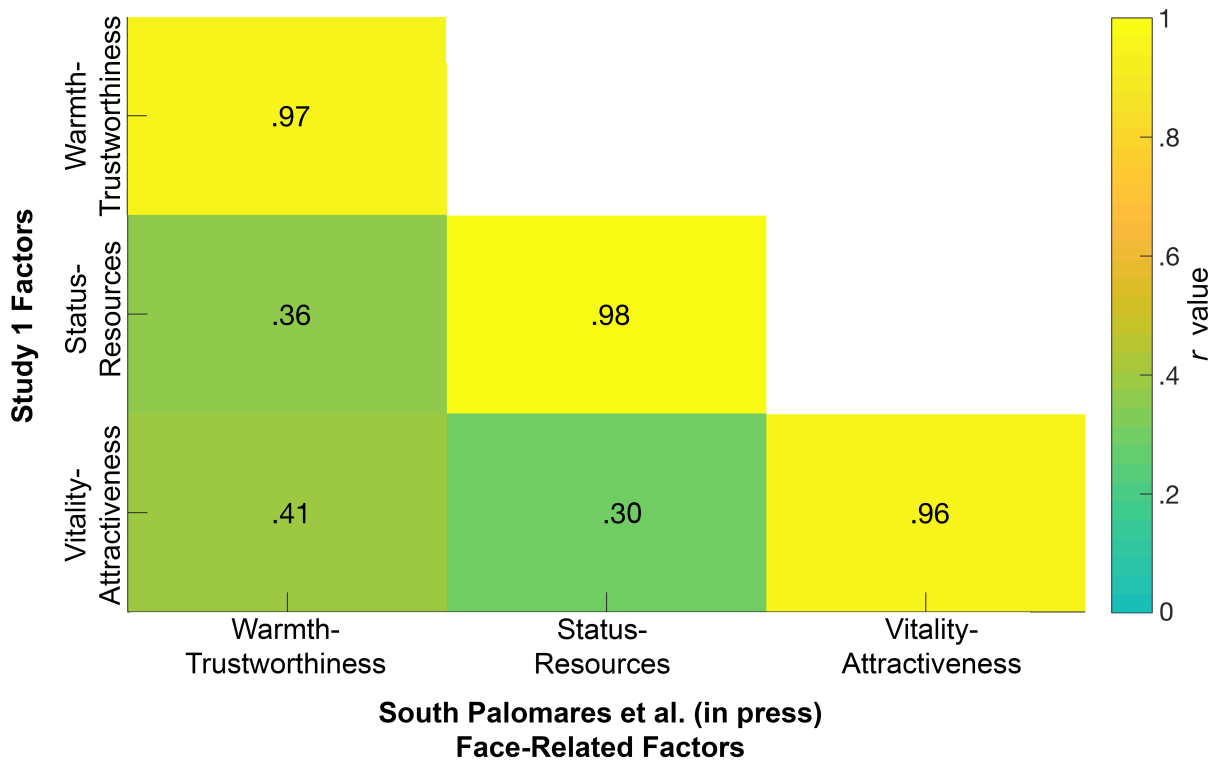


Figure 2. Correlations between the face-related model from South Palomares et al. (2018) and the factor scores for the face-related model derived from potential partner preference-related traits from Fletcher et al. (1999) in Study 1. These are presented in terms of the strength of correlation. All correlations ($N = 1,000$ images in each correlation) were significant at $p < .001$. The strongest correlations are between (1) the warmth-trustworthiness factors; (2) the status-resources factors; and (3) the vitality-attractiveness factors.

Discussion

Study 1 examined the potential of using ambient images as measures of Fletcher et

al.'s (1999) partner preference factors, with 12 traits selected to directly represent the factors Fletcher et al. derived from self-report questionnaires. Findings revealed that all traits could be consistently estimated from face images and a factor analysis of the face ratings revealed a three-factor structure, largely overlapping with Fletcher et al.'s factors. The main difference was that for face images the traits adventurous and vitality loaded strongly on both vitality-attractiveness and warmth-trustworthiness factors. This pattern was also noted by South Palomares et al. (2018), who used a different selection of traits to arrive at a very similar model (see Figure 2). It seems that a subset of facial cues related to adventurousness and vitality overlap with those cues used to represent traits linked to both vitality-attractiveness and warmth-trustworthiness. Yet, when considering adventurous and vitality traits in the abstract (based on a verbal measure), these traits may appear quite distinct from traits related to warmth-trustworthiness. Further research might usefully explore this potentially interesting difference between facially-derived and verbally-derived models. For convenience, we will therefore refer later to the factors derived from Study 1 as face-related partner preference factors.

Study 2 explored how explicit evaluations of partner preferences from ambient face images would relate to the underlying structure and component traits from Study 1, and how these facial preferences would relate to preferences expressed through a standard verbal self-report questionnaire.

Study 2

Study 1 identified a strong correspondence between the underlying structure of traits related to partner preferences seen in ambient images of faces and Fletcher et al.'s (1999) verbal self-report partner preference factors. However, this correspondence was derived from trait-specific ratings obtained in contexts that did not mention partner preferences *per se*. In Study 2, heterosexual participants rated 500 opposite-sex ambient images based on how

interesting they found each face in terms of being their ideal partner, and they also completed Fletcher et al.'s (1999) questionnaire. We examined (1) how the facial partner preferences of individual participants would relate to the factor structure found in Study 1, and (2) how facial partner preferences relate to partner preferences from the questionnaire.

Method

Participants. Fifty participants, university students, were recruited via the University (50% male, mean age of 21 years, $SD = 3.66$). Participants were self-reported heterosexual native English speakers, raised in a Western environment. The participants provided written consent to procedures approved by the Ethics Committee of the University Psychology Department. Participants did not take part in the other reported studies.

Face images. The same database described for Study 1.

Verbal questionnaire measure. The Ideal Partner Scale (Fletcher et al., 1999) is a widely-used measure of verbal self-report partner preferences with good internal reliability, test-retest reliability, and convergent and predictive validity (Fletcher et al., 2000, 1999). Participants rate items on their importance for describing their ideal partner in a close relationship (1: very unimportant – 7: very important): sexy, nice body, attractive appearance, good lover, outgoing, and adventurous (representing vitality-attractiveness); successful, nice house, financially secure, dresses well, and good job (representing status-resources, each accompanied by ‘*or potential to achieve*’ in parentheses); and kind, supportive, understanding, considerate, sensitive, and good listener (representing warmth-trustworthiness).

Procedure. Participants were informed that the study involved romantic facial evaluations and that the task was self-paced, but to rely on their first impressions (South Palomares et al., 2018; South Palomares & Young, 2017; Sutherland et al., 2013; Todorov et al., 2005). Participants rated opposite-sex faces on: “Rate each person on how interesting you

find them in terms of being your ideal partner in a close relationship (e.g., dating, living together, or married)”, using a seven-point Likert scale (1: not interesting – 7: very interesting). Participants completed six practice trials, rating faces randomly selected from the database, and then rated 500 images in a random order. Images remained on the screen until participants had made their judgement and the inter-trial interval was 750 ms. Participants then completed the Ideal Partner Scale (Fletcher et al., 1999). On completion, participants were debriefed and reimbursed with a small payment. The task was programmed using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, USA) and took 30 minutes.

Results

Study 2, firstly, investigated how facial partner preferences would relate to the underlying factor structure and component traits from Study 1. Secondly, Study 2 examined how partner preferences from faces relate to verbally expressed preferences from the questionnaire.

To establish which traits accounted for most of the variance in partner preference ratings of faces, we took the partner preference ratings to each image for each individual participant in Study 2 and correlated these with the overall mean rating achieved by the same images (in the data of Study 1) for each of the 12 traits from Study 1. For example, a participant would have rated 500 images as potential partners in Study 2, and we correlated these 500 ratings with the averaged rating for each image across each of the 12 traits from Study 1, leading to 12 correlations per participant. Our reasoning was that the size of these correlations would reflect the extent to which each trait might be involved in the facially-derived partner preference evaluation. Hence if, for instance, a participant valued warmth-trustworthiness then we would expect substantial correlations with the traits trustworthiness, supportive, understanding, and warmth.

Figure 3 presents these individual participant correlations. The correlations for the

four vitality-attractiveness traits revealed moderate-to-strong relationships between these traits and the partner preference face ratings. In contrast, the correlations for the four status-resources traits and the four warmth-trustworthiness traits revealed only weaker relationships.

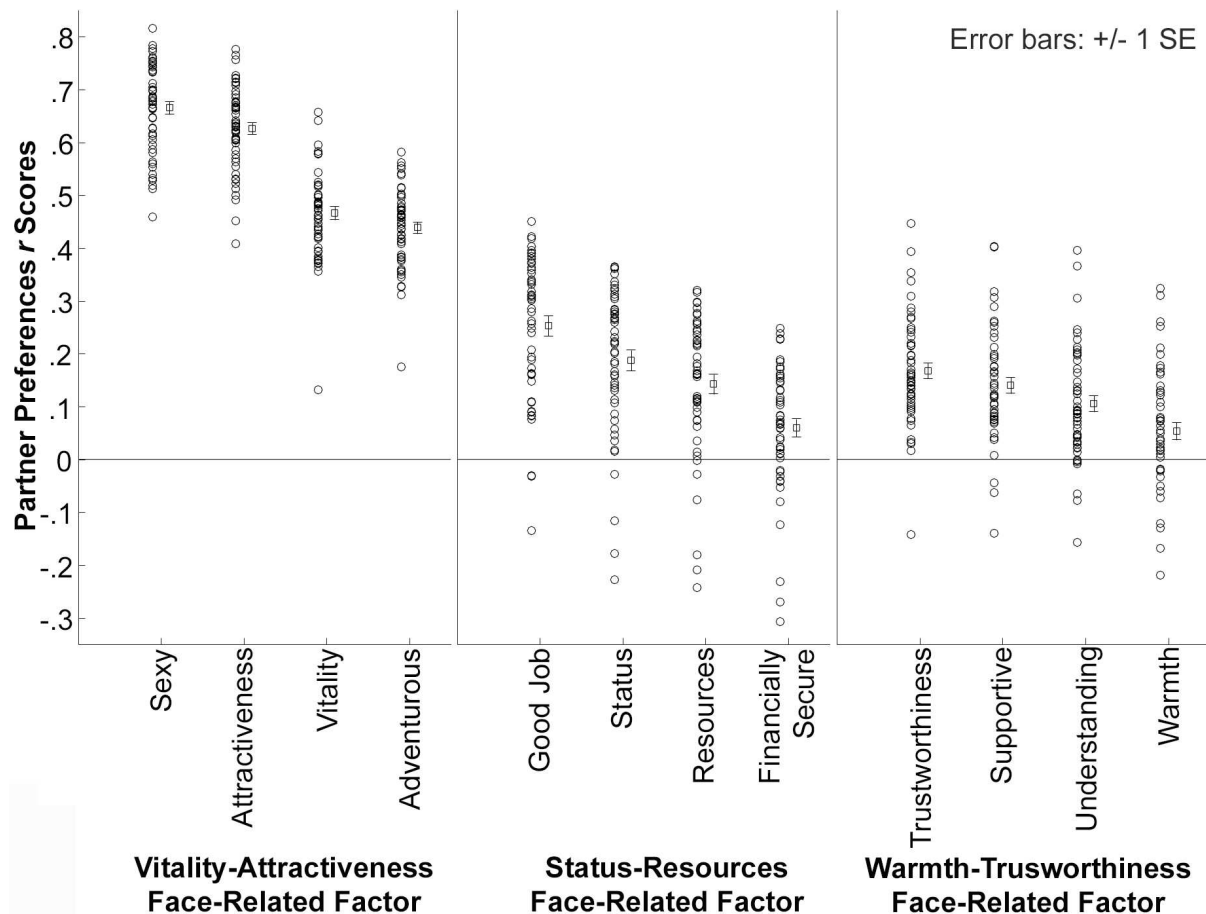


Figure 3. Correlations between individual participant partner preference face ratings to 500 opposite-sex ambient faces in Study 2 and the overall mean trait ratings for the face-related factors derived from Study 1. Higher r values for traits indicate these are prioritised in participants' partner preferences. Circles represent data for individual participants, squares represent mean r values.

The R^2 values of the individual participant correlations across the four traits representing each factor showed that vitality-attractiveness traits were prioritised and

accounted for up to 66% ($M = 32\%$, $SD = 0.02$) of the variance in facial partner preferences. In contrast, the status-resources and warmth-trustworthiness traits respectively accounted for a mean 5% ($SD = .02$, maximum variance = 20%) and 3% ($SD = .01$, maximum variance = 20%) of the variance in facial preferences. Hence, vitality-attractiveness concerns were more salient than status-resources or warmth-trustworthiness concerns when evaluating faces on partner preferences.

To explore the correspondence between facial and verbal questionnaire-based partner preferences, we derived weighted scores for each participant's reliance on the three face-related factors identified in Study 1 in making their facial partner preference judgements. To obtain a weighted score for each participant's preference ratings of the individual faces that could be related to the face-related factors, data from Study 1 were used to compute weighted factor scores representing a loading for each factor in each image. Each participants' partner preference ratings of the face images presented in Study 2 was then correlated with each of these factor scores. Correlations were transformed using Fisher's z (Fisher, 1915) for statistical analyses. They are summarised (as Fisher's z scores) in the upper panel of Figure 4, which shows that the face-related vitality-attractiveness factor accounted for most of the variance in partner preference face ratings. Indeed, every one of our 50 participants prioritised vitality-attractiveness, suggesting it is an overriding concern.

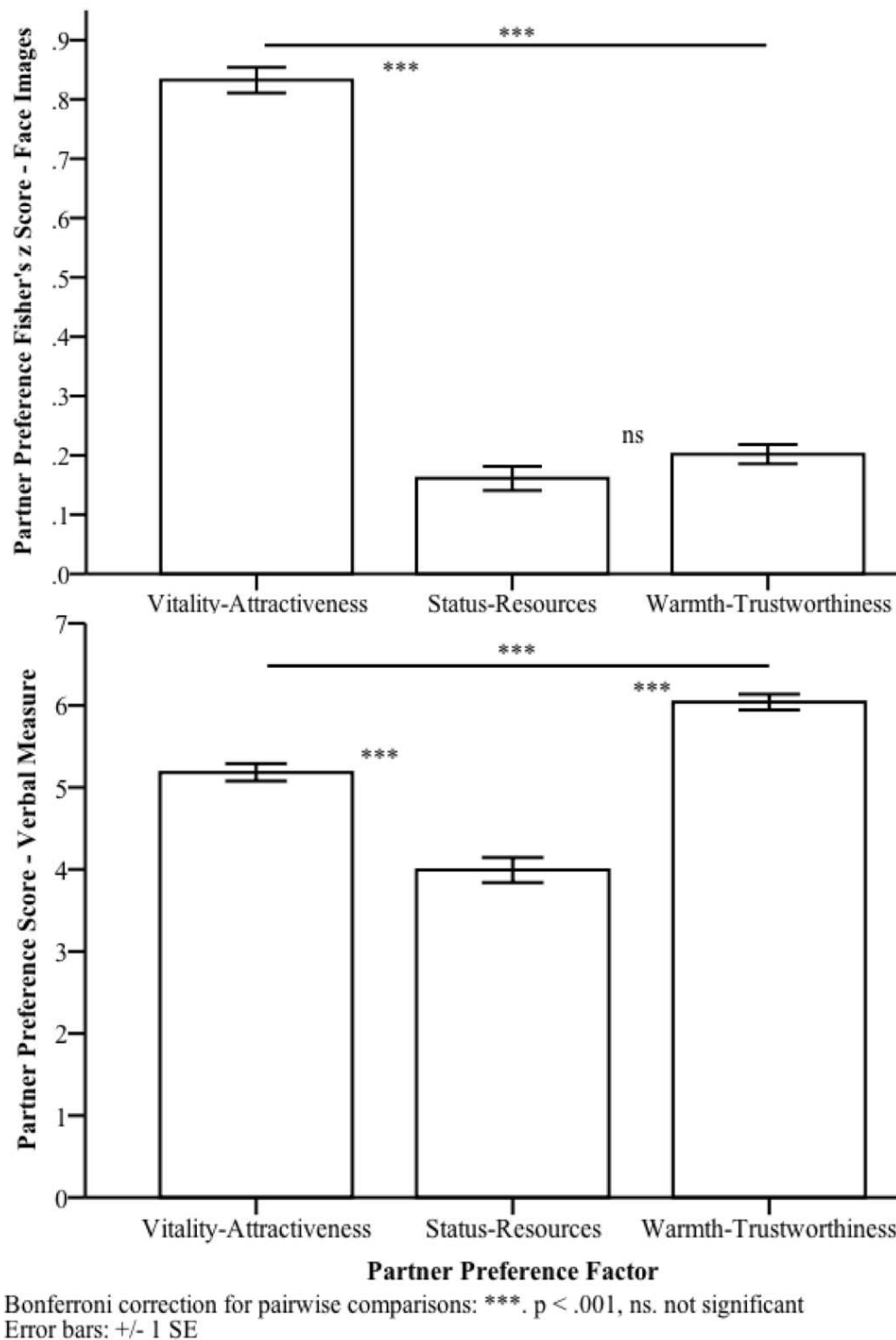


Figure 4. Mean Fisher's z-transformed correlations of partner preferences in Study 2 for ambient faces with face-related factors (top panel), and mean verbally expressed partner preferences from Fletcher et al.'s (1999) questionnaire scale (bottom panel). Larger Fisher's z face scores and larger verbal questionnaire partner preference scores reveal a greater preference for a particular factor.

Regarding the verbal questionnaire preferences, each participant's mean score from the Ideal Partner Scale was calculated for each of the three verbal factors (summarised in the lower panel of Figure 4). To examine the correspondence between facial and verbal questionnaire preferences, the partner preference Fisher's z scores for faces and the questionnaire preference scores were subjected to a repeated-measures 3 x 2 ANOVA between the partner preference factors (vitality-attractiveness, status-resources, and warmth-trustworthiness) and stimuli (face images, verbal questionnaire measure). The model did not violate sphericity. The ANOVA revealed a significant interaction effect between factors and stimuli ($F(2, 98) = 82.93, MSE = 27.13, p < .001, \eta_p^2 = .63$). This interaction represents the key finding as it reveals differences between the patterns of preferences expressed to faces and the verbal questionnaire. There was also a significant main effect of factors ($F(2, 98) = 100.29, MSE = 32.85, p < .001, \eta_p^2 = .67$) and a main effect of stimuli ($F(1, 49) = 3692.50, MSE = 1637.96, p < .001, \eta_p^2 = .99$) that simply reflects the difference between the ranges of the facial and verbal questionnaire scales.

Examination of the critical interaction between partner preference factors and stimuli using Bonferroni's adjustment for multiple pairwise comparisons revealed that, for preferences measured using faces, participants prioritised vitality-attractiveness relative to status-resources and warmth-trustworthiness ($ps < .001$, see Figure 4). There was no significant difference between the face-related status-resources and warmth-trustworthiness factors ($p = .223$). For verbal questionnaire preferences, participants prioritised warmth-trustworthiness, followed by vitality-attractiveness, with status-resources ranking last. There were significant differences ($ps < .001$) between each of the verbal questionnaire factors.

Discussion

Study 2 examined (1) how explicit facial partner preferences relate to the component traits and factors from Study 1, and (2) the relationship between preferences measured to

faces and the verbal self-report questionnaire. Findings revealed that vitality-attractiveness concerns dominated both male and female participants' facial partner preferences, yet, warmth-trustworthiness was prioritised in participants' verbal questionnaire preferences. The questionnaire findings correspond with previous self-report studies (e.g., Buss, 1989; Fletcher et al., 1999, 2004; Li et al., 2002), where participants also typically prioritise warmth-trustworthiness. Hence, it can be inferred that face stimuli and verbal questionnaire self-report measures capture different elements of preferences.

Importantly, physical attractiveness is not necessarily the strongest social signal offered by faces. Instead, valence/warmth/trustworthiness forms the most salient dimension in many facial first impressions studies (including Study 1 here), not attractiveness (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Vernon et al., 2014). We will discuss later possible reasons why people may prioritise facial attractiveness for potential partners.

Study 2 used adult faces unconstrained on age, but many studies link attractiveness to age (Henss, 1991; Jones & Hill, 1993; Matts, Fink, Grammer, & Burquest, 2007; Sutherland et al., 2013; Tatarunaite, Playle, Hood, Shaw, & Richmond, 2005). Hence, given our young adult participants (mean age: 21 years), we thought it necessary to test whether the emphasis on vitality-attractiveness that we found may have been a by-product of their ruling out the older-looking faces. Some research also suggests that younger participants do not discriminate well between the ages of faces older than themselves (van Rijsbergen, Jaworska, Rousselet, & Schyns, 2014), underscoring the need to test in Study 3 whether young adults still prioritise vitality-attractiveness when evaluating relatively youthful-looking faces.

Study 3

Study 2 revealed that when rating faces unconstrained on age all participants prioritised vitality-attractiveness in their partner preference face ratings whereas, in contrast, their verbal questionnaire preferences prioritised warmth-trustworthiness. Study 3 examined

whether the patterns of facial preferences of our young adult participants would change if the stimuli were themselves more youthful in appearance. Although we had a sizeable sample of 500 ambient images of each sex in our complete set of stimuli, the constraints of using only the more youthful-looking opposite-sex images would have reduced the numbers of stimuli to the point where we were concerned that the potential advantages of the ambient image method might be lost. We therefore used image-averaging techniques to create relatively youthful face averages from images scoring high or low on vitality-attractiveness, status-resources, and warmth-trustworthiness.

Averaging sets of ambient images can create face-like prototype images that maintain the consistent cues governing the perception of particular traits (South Palomares et al., 2018; South Palomares & Young, 2017; Sutherland et al., 2013; Sutherland, Rhodes, et al., 2017). We used the same tactic to create averages of 10 ambient images of young adults with high or low weightings on the factor in question, which were then evaluated on partner preferences by a new group of student participants. Hence, our interest was in whether participants would still prioritise the face-related vitality-attractiveness partner preference factor relative to status-resources and warmth-trustworthiness. As in Study 2, we also examined the correspondence between facial and verbal questionnaire preferences.

Method

Participants. Forty participants, university students, were recruited via the University (50% male, mean age of 21 years, $SD = 4.00$). Participants were self-reported heterosexual native English speakers, raised in a Western environment. The participants provided written consent to procedures approved by the Ethics Committee of the University Psychology Department. Participants did not take part in the other reported studies.

Face-like images. Using the factor loadings for each image from Study 1, we followed a previously adopted procedure (South Palomares et al., 2018; South Palomares &

Young, 2017; Sutherland et al., 2013) to create face-like averaged images representing high and low levels of each factor using PsychoMorph (Tiddeman, Burt, & Perrett, 2001). Each image in Figure 5 was made by averaging the 10 highest and lowest loading images on the relevant factor, then the 10 next highest and 10 next lowest loading images, and so on. Only images representing male or female younger adults were selected, based on age ratings from Santos & Young (2005, 2008, 2011) in which participants rated the 1,000 face stimuli on age using a Likert scale (*1: young adult to 7: old adult*). Stimuli with average ratings less than or equal to 4.00 were selected to ensure that the images represented young adults.

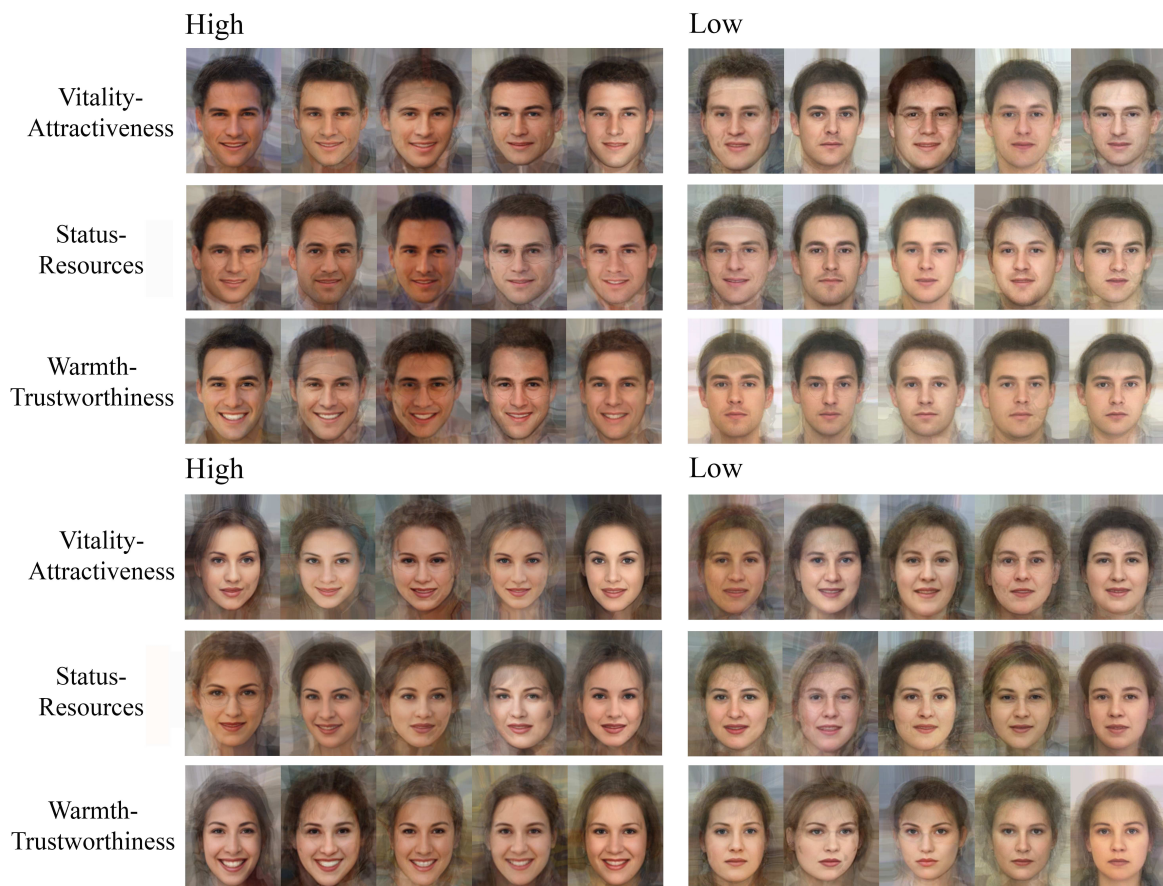


Figure 5. Stimuli used in Study 3. These are averaged youthful-looking face images. Each average is made from 10 ambient images representing relatively high (left panels) and relatively low (right panels) levels of the face-related factors identified in Study 1. Averages created from male images are shown in the upper panels and female images in the lower panels.

Verbal questionnaire measure. The study used the same Ideal Partner Scale described for Study 2.

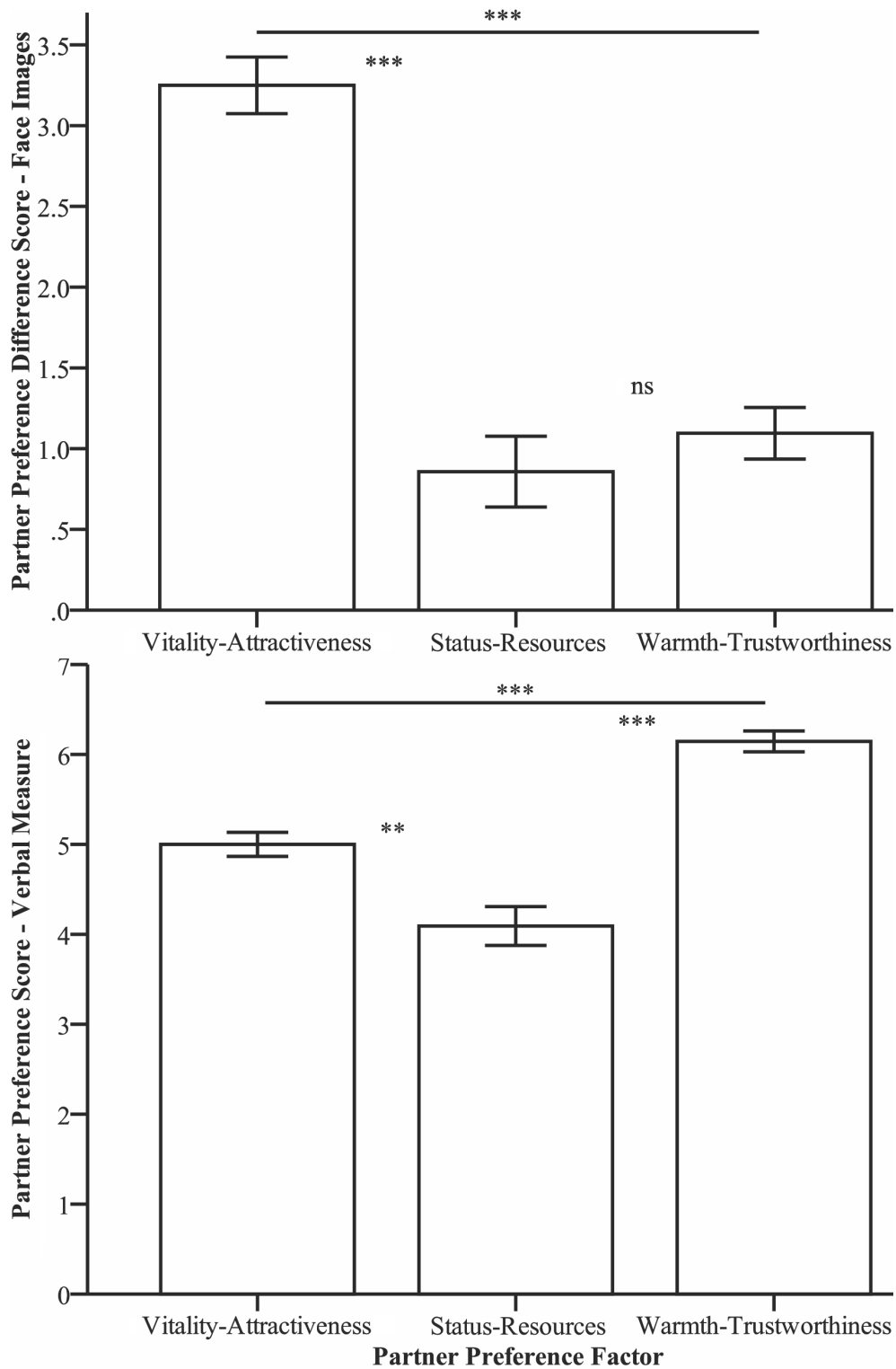
Procedure. All procedural details were the same as Study 2, excepting that participants rated partner preferences for 30 opposite-sex face-like averaged images twice in a random order. The task took 10 minutes.

Results

Because Study 3 used averaged images we needed to use a different measure from

Studies 1 and 2. To obtain a score for each participant's ratings of the face-like images in terms of the three face-related partner preference factors, a difference score was calculated based on the individual participant's mean difference in their partner preference ratings for the high and low face-like average images relating to each factor. For example, for the vitality-attractiveness images, the mean rating for the five low vitality-attractiveness averaged images was subtracted from the mean rating for the five high vitality-attractiveness images. This mean rating difference score was calculated separately per factor and per participant. Higher difference scores reflect more emphasis of a given factor in the rated partner preferences. For instance, if Peter rates high vitality-attractiveness faces as very interesting in terms of being his ideal partner (7: very interesting) and Peter is not interested in images low in vitality-attractiveness (1: not interesting), Peter's difference score would be the maximum possible of +6, reflecting a strong preference for vitality-attractiveness.

Using this measure all participants, except one, prioritised vitality-attractiveness (vitality-attractiveness mean: 3.25; status resources mean: 0.86; warmth-trustworthiness mean: 1.10) for faces. Concerning the verbal questionnaire preferences, individual participant mean scores from the Ideal Partner Scale were calculated for each factor, as in Study 2: vitality-attractiveness mean: 5.00; status resources mean: 4.09; warmth-trustworthiness mean: 6.15). Figure 6 summarises data for the facial and verbal questionnaire partner preference measures.



Bonferroni adjustment for pairwise comparisons: ***. $p < .001$, **. $p < .01$, n.s. not significant. Error bars: ± 1 SE

Figure 6. Partner preferences for youthful adult face-like averaged images (top panel) and verbal questionnaire partner preferences (bottom panel) in Study 3. Larger face difference scores and higher verbal questionnaire preference scores reveal a greater preference for a particular factor.

To explore the correspondence between facially derived and verbally expressed partner preferences, individual participants' mean face difference scores and their verbal questionnaire scores were subjected to a repeated-measures 3 x 2 ANOVA of the partner preference factors (vitality-attractiveness, status-resources, and warmth-trustworthiness) and stimuli (face images, verbal questionnaire measure). A Greenhouse-Geisser correction was used as the model violated sphericity. This ANOVA revealed a significant interaction effect between factors and stimuli ($F(1.73, 67.60) = 46.00, MSE = 63.07, p < .001, \eta_p^2 = .54$). This interaction represents the key finding, as it reveals differences between the patterns of partner preferences expressed to facial and verbal questionnaire stimuli. Despite the change in stimuli and measure for the face-based part of Study 3, the overall pattern of findings closely resembled those from Study 2 (see Figure 6). There was, also, a significant main effect of factor ($F(1.91, 74.41) = 68.85, MSE = 59.95, p < .001, \eta_p^2 = .64$) and a significant main effect of stimuli ($F(1.00, 39.00) = 323.56, MSE = 671.45, p < .001, \eta_p^2 = .89$). Examination of the critical interaction between partner preference factors and stimuli using Bonferroni's adjustment revealed similar results to Study 2: for preferences measured using faces, participants prioritised vitality-attractiveness relative to status-resources and warmth-trustworthiness ($ps < .001$, see Figure 6). There was no significant difference between the face-related status-resources and warmth-trustworthiness factors ($p = .79$). For verbal questionnaire preferences, participants prioritised warmth-trustworthiness, followed by vitality-attractiveness, with status-resources ranking last. There were significant differences ($ps < .01$) between each of the verbal questionnaire factors.

Discussion

Study 3 investigated partner preferences for relatively youthful-looking averaged

face-like images and how these relate to preferences from the verbal questionnaire. As for Study 2, facial and verbal questionnaire preferences were not comparable. Participants strongly prioritised vitality-attractiveness in measures of partner preferences using these more youthful-looking face-like stimuli, yet, they prioritised warmth-trustworthiness in their verbal preferences. We will discuss later possible reasons for the disparity between facial and verbal questionnaire preferences. The fact that this research replicated the pattern of findings using different stimuli (Study 2: ambient images; Study 3: averaged images) lends confidence in the generality of the effect.

Given the consistency of the dominance of physical attractiveness in partner preference ratings to faces found in Studies 2 and 3, we were interested in whether this is an inevitable concomitant of judging partner preferences, or whether more nuanced evaluations might be possible if other partner preference-related factors are foregrounded by the task instructions. Study 4, therefore, explored whether potential partner preferences for faces are dominated by physical attractiveness even when other potentially important factors are made salient.

Study 4

Studies 2 and 3 found that participants prioritise vitality-attractiveness when rating ambient face images and youthful-looking averaged faces as potential partners. Study 4 examined whether the factors influencing potential partner preferences to faces can be changed by asking participants to focus on a particular relationship type. Participants rated opposite-sex faces on their partner preferences for a specific relationship type. We were interested in establishing whether facial preferences are sufficiently flexible that they can shift when different specific relationship types are made salient. Relationship descriptors were selected based on a preliminary study: exciting for a relationship based on vitality-attractiveness, lavish for a relationship needing status-resources, and loyal for a relationship

grounded in warmth-trustworthiness.

Method

Participants. Sixty participants, university students, were recruited via the University (50% male, mean age of 21 years, $SD = 3.02$). Participants were self-reported heterosexual native English speakers, raised in a Western environment. The participants provided written consent to procedures approved by the Ethics Committee of the University Psychology Department. Participants did not take part in the other reported studies.

Relationship descriptors. Three relationship descriptors were chosen to represent the partner preference factors based on a preliminary study in which participants rated descriptors we thought might represent each factor (vitality-attractiveness descriptors: exciting, fun, and passionate; status-resources descriptors: lavish, extravagant, and luxurious; and warmth-trustworthiness descriptors: loyal, intimate, and committed). One descriptor was selected to represent each factor based on a repeated-measures ANOVA between factor and descriptor. For instance, the loyal descriptor obtained significantly higher ratings on a warmth-trustworthiness relationship relative to its ratings on vitality-attractiveness or status-resources relationships. The selected descriptors were: vitality-attractiveness - exciting; status-resources - lavish; and warmth-trustworthiness - loyal.

Face images. The same database described for Study 1.

Procedure. Participants rated each of 500 opposite-sex images for "how interesting you find them in terms of being your ideal partner in a LOYAL close relationship (e.g., dating, living together, or married)". The word "loyal" was replaced with the relevant relationship descriptor (exciting, lavish, or loyal) as appropriate. All other procedural details were the same as for Study 2, except that groups of 20 participants (10 male) were randomly assigned to rate images based on one of the three relationship types to avoid carry-over effects (Rhodes, 2006).

Results

This study investigated how facial partner preferences as a function of relationship type relate to the underlying factor structure and component traits from Study 1. As per Study 2, we took the partner preference ratings to each of 500 images for each participant, separated by relationship type, and correlated these with the overall mean rating achieved by the same images (data from Study 1) for each of the 12 traits used in Study 1.

Figure 7 presents these individual participant correlations for each relationship type. Independent of relationship type, the correlations for the four vitality-attractiveness traits revealed moderate-to-strong relationships between these traits and the partner preference face ratings. In contrast, the four status-resources traits and the four warmth-trustworthiness traits revealed only weak relationships. Almost every one of the 60 participants prioritised the face-related vitality-attractiveness factor, regardless of relationship type, indicating once more that vitality-attractiveness was largely an overriding concern.

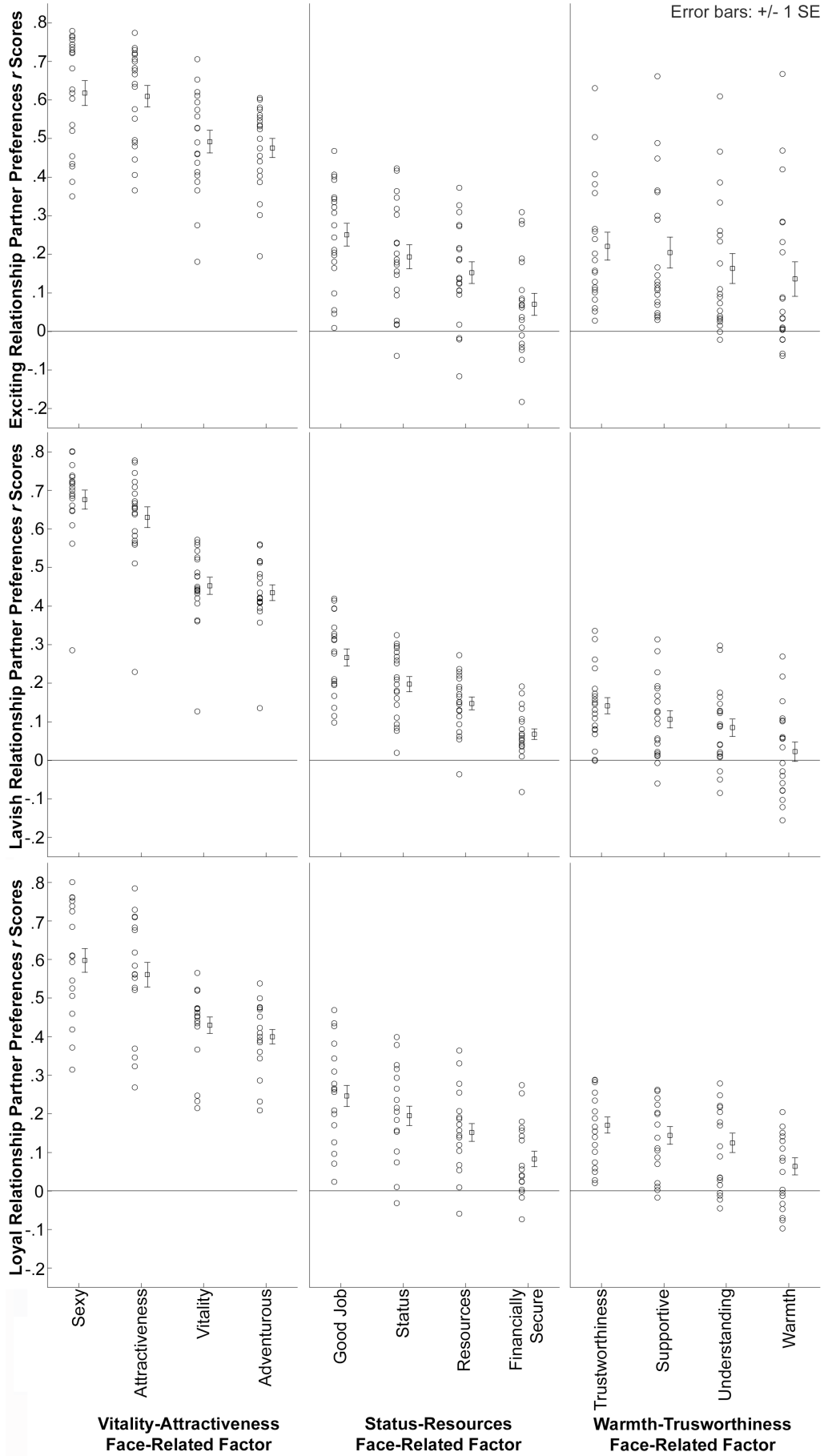
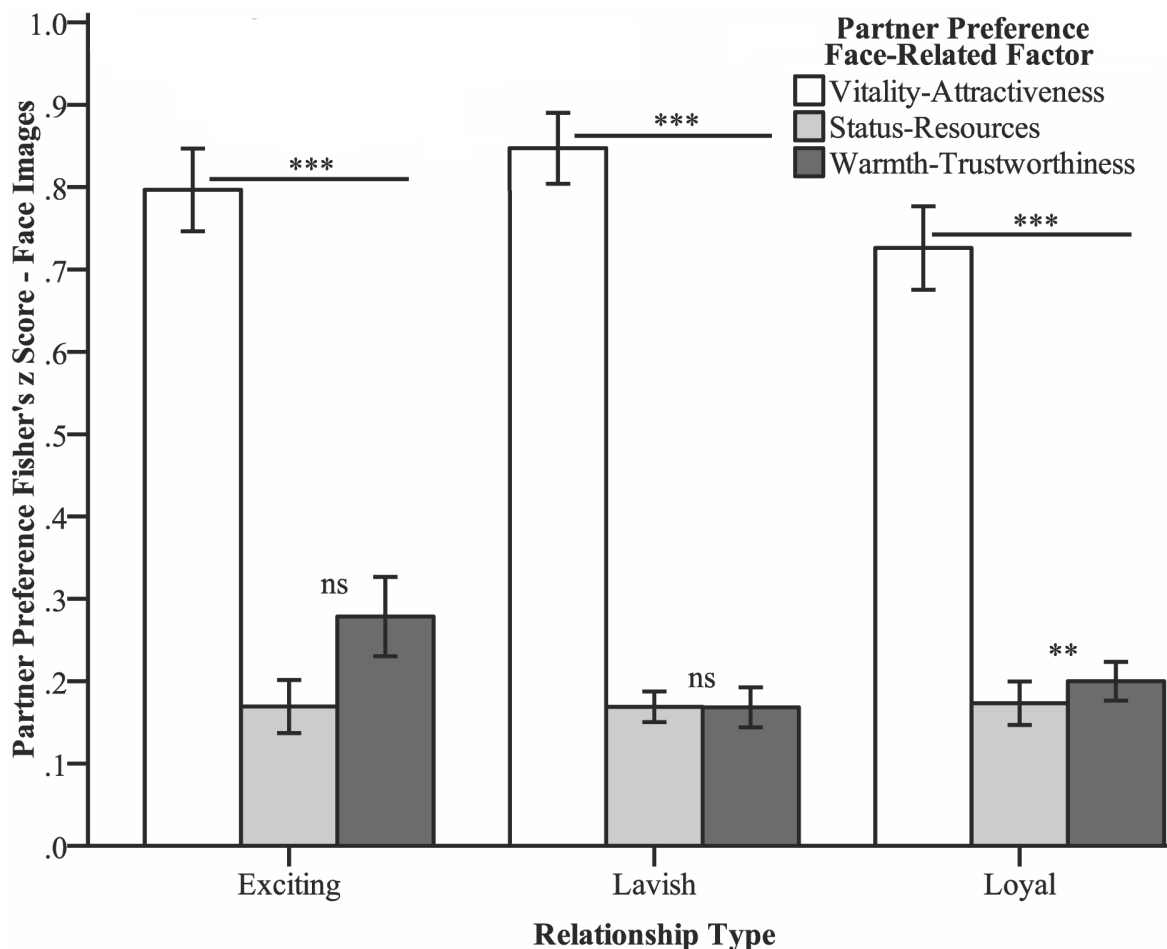


Figure 7. Correlations between individual participant partner preference face ratings for an exciting relationship (top panel), a lavish relationship (middle panel), and a loyal relationship (bottom panel) to opposite-sex ambient face images in Study 4 and the overall mean trait ratings for the face-related factors derived from Study 1. Higher r values for traits indicate these are prioritised in participants' partner preferences. Circles represent data for individual participants, squares represent mean r values.

The R^2 values of the individual participant correlations across the four traits representing each face-related factor, separated by relationship type, revealed that vitality-attractiveness traits were prioritised and overall accounted for up to 64% ($M = 30\%$, $SD = 0.15$) of the variance in facial partner preferences, independent of relationship type. In contrast, the status-resources and warmth-trustworthiness traits respectively accounted for a mean 4% ($SD = .05$, maximum variance = 22%) and 4% ($SD = .07$, maximum variance = 44%) of the variance in facial preferences. Therefore, vitality-attractiveness dominated facial partner preferences, independent of relationship type.

To examine facial partner preferences in terms of relationship type using the face-related factors derived from Study 1, we again calculated weighted scores for each participant's reliance on the three face-related factors identified in Study 1 and then correlated participants' partner preference face ratings from Study 4 (for their assigned relationship) with these factor scores. The resulting partner preference correlations were transformed using Fisher's z (Fisher, 1915) and subjected to a mixed 3 x 3 ANOVA between relationship type (exciting, lavish, and loyal) and partner preference face-related factor (vitality-attractiveness, status-resources, and warmth-trustworthiness). A Greenhouse-Geisser correction was used as the model violated sphericity. Data are presented in Figure 8. There was no significant interaction effect between relationship and factor ($F(2.96, 84.24) =$

2.02, $MSE = 0.07$, $p = .118$, $\eta_p^2 = .07$). This lack of an interaction represents the key finding, revealing that participants did not shift their partner preferences in accordance with the type of relationship specified in the task instructions. Findings, further, revealed a significant main effect for factor ($F(1.48, 84.24) = 281.20$, $MSE = 9.69$, $p < .001$, $\eta_p^2 = .83$), but not for relationship ($F(2, 57) = 1.10$, $MSE = 0.04$, $p = .340$, $\eta_p^2 = .04$). Pairwise comparisons using Bonferroni's adjustment revealed that participants prioritised vitality-attractiveness relative to status-resources and warmth-trustworthiness ($ps < .001$) with no significant difference between the status-resources and warmth-trustworthiness factors ($p = .089$). These findings are consistent with findings from Studies 2 and 3, revealing that individuals prioritise vitality-attractiveness.



Bonferroni adjustment for pairwise comparisons: ***, $p < .001$, **, $p < .01$, ns. not significant
 Error bars: ± 1 SE

Figure 8. Mean Fisher's z-transformed correlations of partner preferences in terms of relationship type in Study 4 for ambient faces, as a function of the face-related factors from Study 1. Larger Fisher's z face scores reveal a greater preference for a particular factor.

Discussion

Study 4 examined the malleability of partner preferences to faces as a function of relationship type (exciting, lavish, or loyal, chosen to emphasise the different factors of vitality-attractiveness, status-resources, and warmth-trustworthiness respectively). Regardless of relationship type, participants prioritised vitality-attractiveness in their facial preferences. Hence, prioritising attractiveness in a first impression of a potential partner's face may be a difficult tendency to overcome. Studies 2 to 4 found little influence of the status-resources

and warmth-trustworthiness face-related partner preference factors we had identified in Study 1, with all facial preference judgements being dominated by vitality-attractiveness.

General Discussion

Study 1 showed that it might in principle be possible to use responses to faces as measures of Fletcher et al.'s (1999) verbal questionnaire self-report partner preference factors by demonstrating that the traits and factors identified by Fletcher et al. (1999) can also be seen in images of faces, with a comparable underlying three-factor structure. However, when Studies 2 and 3 evaluated how explicit judgements of facial partner preferences actually relate to verbally expressed preferences from questionnaire responses, it was clear that they did not correspond. Instead, physical attractiveness traits were found to dominate the evaluations of faces for nearly all participants, with little role for the other traits or factors. Study 4 investigated whether it might be possible for the different factors to be used more flexibly in facial partner preferences, by asking participants to focus on a type of relationship that was linked to a specific factor, but again facial attractiveness dominated their evaluations.

This dominance of facial attractiveness led to poor correspondence between facial and verbal questionnaire measures of partner preference. Verbally expressed self-reported preferences measured by Fletcher et al.'s (1999) questionnaire involved a relatively balanced mix of evaluations involving all three factors, but with a tendency to emphasise warmth-trustworthiness (see Figures 4 and 6). This pattern contrasted strikingly with the consistently strong emphasis on vitality-attractiveness in evaluating faces (Figures 3, 4, 6, 7, and 8).

This disparity in the role of physical attractiveness between verbally and facially-derived partner evaluations needs to be explained. For verbal preferences, our findings correspond with previous questionnaire-based/self-report research (e.g., Buss, 1989; Fletcher et al., 1999, 2004; Li et al., 2002), where participants typically prioritise warmth-

trustworthiness, then vitality-attractiveness and status-resources. Thus, facial and verbal self-report measures may capture different elements of preferences.

Quick evaluations of everyday faces image are frequent in online contexts; for this reason, our approach was grounded in facial first impressions research. This research reveals that, as we also found in Study 1, a combination of warmth/trustworthiness/approachability/valence emerges as the factor that generally accounts for the largest proportion of the variance in facial first impressions (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Vernon et al., 2014), and that the evaluation of warmth-related traits is easily achieved from even a brief glimpse of a face (South Palomares & Young, 2017; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Yet, the present findings revealed that it was vitality-attractiveness (rather than warmth-trustworthiness) which dominated facial partner preferences. The extent to which vitality-attractiveness is prioritised in partner preferences above other traits that are easily seen in faces is, therefore, an interesting and important phenomenon.

Two broad theoretical approaches may be relevant to understanding the salience of facial attractiveness in partner preferences. The first approach derives from Dion, Berscheid, and Walster's (1972) observations of a halo effect of attractiveness on a wide range of social judgements. Notably, relative to less attractive people, attractive people are perceived to be better romantic partners (Dion et al., 1972) and to have better job prospects (Chiu & Babcock, 2002; Marlowe, Schneider, & Nelson, 1996), showing that the attractiveness halo can permeate more than one of the Fletcher et al. (1999) factors. However, in other contexts where facial attractiveness can have a clear influence, such as electoral success (Berggren, Jordahl, & Poutvaara, 2010), other factors besides attractiveness also show substantial influences (Chen, Jing, & Lee, 2014; Todorov et al., 2005). Similarly, other work using the same stimuli as our experiments has shown that facial evaluations in the non-romantic

context of judging occupations draw flexibly on different traits in ways which reflect the stereotypical attributes associated with particular social groups (Oldmeadow, Sutherland, & Young, 2013). This was not the case here. Instead, nearly all participants in Studies 2 to 4 relied heavily on vitality-attractiveness in evaluating partner preferences from faces, with the other factors of status-resources and warmth-trustworthiness accounting for little of the variance in responses. Taking these findings together, it is clear that the emphasis on vitality-attractiveness in potential partner preferences to faces is not due to vitality-attractiveness being intrinsically more salient, relative to other traits or factors, in any face evaluation context. Rather, the findings underscore the point that, whilst vitality-attractiveness dominates romantic evaluations of faces, it does not dominate many non-romantic evaluations.

The second theoretical approach links physical attractiveness to mechanisms of sexual selection (Fink & Penton-Voak, 2002; Rhodes, 2006). For example, Symons (1979) compared attractiveness to a 'health certificate' providing cues to fertility and resistance to environmental and genetic stressors (Jasienska, Lipson, Ellison, Thune, & Ziomkiewicz, 2006; Lassek & Gaulin, 2008; Møller, 1999; see Rhodes, 2006, for a review). Nonetheless, it is clear that some aspects of attractiveness are learnt (Dotsch, Hassin, & Todorov, 2016; Perrett et al., 2002) and these learnt preferences can lead to inter-individual differences (Germine et al., 2015). It thus seems likely that multiple influences are operating (Maestripieri, Henry, & Nickels, 2017).

Within the romantic domain, then, there are a number of reasons why attractiveness-related traits may be differentially salient on first acquaintance, compared to benefits derived from prioritising traits linked to status-resources or warmth-trustworthiness. In contrast, studies of verbally expressed self-reports of partner preferences emphasise that other factors

including cultural influences and personal priorities come into play (Zentner & Mitura, 2012).

Our studies applied a relatively novel measure of partner preferences offering ecological validity through using ambient images, which can capture the role of cues that might be absent in standardised images (Burton, Jenkins, & Schweinberger, 2011; Jenkins et al., 2011). Nonetheless, standardised images remain useful in improving sensitivity to detect small effects and in manipulating images to determine causal effects. A strength of the present research is the complementary use of ambient and averaged face-like images to show that multiple interacting cues may contribute to partner preferences. In Studies 2 and 4 ambient stimuli allowed these cues to vary naturally, facilitating the identification of the main traits involved in partner preferences, which is a theoretically powerful approach (Sutherland et al., 2013; Vernon et al., 2014; Young, 2018). In study 3, averaged images were used to emphasise the characteristics common to a specific factor. The complementary use of both ambient and standardised images offers a more comprehensive understanding of person perception (Sutherland, Rhodes, et al., 2017).

Whilst our use of everyday face images closely approximates what happens in contexts such as online dating and relationships, these contexts are clearly impoverished in comparison to face to face encounters. However, other studies have also found that attractiveness is prioritised in relatively realistic contexts (Walster, Aronson, Abrahams, & Rottman, 1966). Todd et al. (2007, p. 15015) used a speed-dating paradigm to examine men's partner preferences and suggested these "are best described as an avoidance of unattractiveness". Fletcher et al. (2014) asked opposite-sex strangers to engage in a 10-minute date; they found that vitality-attractiveness was a more salient and a more accurate predictor of romantic interest than status-resources or warmth-trustworthiness. Eastwick, Luchies, Finkel, and Hunt (2014) revealed that attractiveness predicted partner preferences

with a moderate-to-strong effect size whereas earning prospects predicted preferences with a small effect size. Hence, the difference between verbal self-reports and facial measures of partner preferences may be linked to differences in context and level of processing (e.g., abstract and conceptual verbal statements versus more perceptually-driven responses towards specific face images), amongst other possible factors.

Our findings speak directly to the theoretical question of the flexibility of person construal. Hassin and Trope (2000) and Macrae and Bodenhausen (2000) emphasised the issue of whether perceivers can control the influences of different types of social evaluation. Subsequent studies have shown some flexibility to this process (Quadflieg et al., 2011; Quinn & Macrae, 2005; Santos & Young, 2005), but our findings suggest the extent of this flexibility is task-dependent. When participants rated faces in terms of partner preferences, they seemed unable to consider traits beyond vitality-attractiveness, despite their more complex verbally expressed preferences. Yet, when judging faces in terms of occupational stereotypes, participants can readily use a wide range of evaluative cues (cf. Oldmeadow et al., 2013). The remarkable differences in performance between these tasks may offer a useful way of approaching questions concerning the factors that govern the flexibility of these pervasive social evaluations, and mechanisms of cognitive control more generally.

To conclude, many relationships begin in contexts where facial impressions will form an important source of information (e.g., online dating). Yet, examining first impressions of potential partners using naturally occurring images is a relatively novel approach, with relationship researchers having primarily used measures based on verbal self-reports. Our studies show that although the traits important in theories of partner preferences derived from verbal responses can also be seen in faces, participants' partner preferences to faces are largely dominated by attractiveness-related concerns. This dominance of facial attractiveness arises even in contexts where it would be important to consider other attributes. Facial and

verbal questionnaire preferences do not fully correspond, then, suggesting that these measures capture different elements of preferences. The prevalence of online images and internet-based relationships make these findings timely and relevant to contexts within and beyond the romantic domain.

References

- Berggren, N., Jordahl, H., & Poutvaara, P. (2010). The looks of a winner: Beauty and electoral success. *Journal of Public Economics*, *94*(1–2), 8–15.
<https://doi.org/10.1016/j.jpubeco.2009.11.002>
- Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology*, *102*(4), 943–958.
<https://doi.org/10.1111/j.2044-8295.2011.02039.x>
- Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*(1), 1–49.
<https://doi.org/10.1017/S0140525X00023992>
- Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, *50*(3), 559–570. <https://doi.org/10.1037//0022-3514.50.3.559>
- Campbell, L., Chin, K., & Stanton, S. C. E. (2016). Initial evidence that individuals form new relationships with partners that more closely match their ideal preferences. *Collabra*, *2*(1), 2. <https://doi.org/10.1525/collabra.24>
- Campbell, L., Simpson, J. A., Kashy, D. A., & Fletcher, G. J. O. (2001). Ideal standards, the self, and flexibility of ideals in close relationships. *Personality and Social Psychology Bulletin*, *27*(4), 447–462. <https://doi.org/10.1177/0146167201274006>
- Chen, F. F., Jing, Y., & Lee, J. M. (2014). The looks of a leader: Competent and trustworthy, but not dominant. *Journal of Experimental Social Psychology*, *51*, 27–33.

<https://doi.org/10.1016/j.jesp.2013.10.008>

- Chiu, R. K., & Babcock, R. D. (2002). The relative importance of facial attractiveness and gender in Hong Kong selection decisions. *International Journal of Human Resource Management, 13*(1), 141–155. <https://doi.org/10.1080/09585190110092857>
- Dion, K. K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285–290. <https://doi.org/10.1037/h0033731>
- Dotsch, R., Hassin, R. R., & Todorov, A. (2016). Statistical learning shapes face evaluation. *Nature Human Behaviour, 1*(1), 1–6. <https://doi.org/10.1038/s41562-016-0001>
- Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin, 140*(3), 623–665. <https://doi.org/10.1037/a0032432>
- eHarmony. (2015). eHarmony UK Statistics: Almost 2.5 million members. Retrieved July 19, 2015, from <http://www.eharmony.co.uk/local-dating/member-index/>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11*(2000), 154–158. <https://doi.org/10.1111/1467-8721.00190>
- Fiore, A. T., Taylor, L. S., Mendelsohn, G. a., & Hearst, M. (2008). Assessing attractiveness in online dating profiles. *Proceedings of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems CHI 08, (1992)*, 797–806. <https://doi.org/10.1145/1357054.1357181>
- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10*(4), 507–521.

<https://doi.org/10.2307/2331838>

Fletcher, G. J. O., Kerr, P. S. G., Li, N. P., & Valentine, K. A. (2014). Predicting romantic interest and decisions in the very early stages of mate selection: Standards, accuracy, and sex differences. *Personality and Social Psychology Bulletin*, *40*, 540–50.

<https://doi.org/10.1177/0146167213519481>

Fletcher, G. J. O., & Simpson, J. A. (2000). Ideal standards in close relationships: Their structure and functions. *Current Directions in Psychological Science*, *9*(3), 102–105.

<https://doi.org/10.1111/1467-8721.00070>

Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). Ideals, perceptions, and evaluations in early relationship development. *Journal of Personality and Social Psychology*, *79*(6), 933–940. <https://doi.org/10.1037//0022-3514.79.6.933>

Fletcher, G. J. O., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology*, *76*(1), 72–89.

<https://doi.org/10.1037/0022-3514.76.1.72>

Fletcher, G. J. O., Tither, J. M., O’Loughlin, C., Friesen, M. D., & Overall, N. C. (2004). Warm and homely or cold and beautiful? Sex differences in trading off traits in mate selection. *Personality and Social Psychology Bulletin*, *30*(6), 659–72.

<https://doi.org/10.1177/0146167203262847>

Germine, L., Russell, R., Bronstad, P. M., Blokland, G. A. M., Smoller, J. W., Kwok, H., ... Wilmer, J. B. (2015). Individual aesthetic preferences for faces are shaped mostly by environments, not genes. *Current Biology*, *25*(20), 2684–2689.

<https://doi.org/10.1016/j.cub.2015.08.048>

Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication*, *59*(2), 367–386.

<https://doi.org/10.1111/j.1460-2466.2009.01420.x>

- Hassin, R. R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, 78(5), 837–852.
<https://doi.org/10.1037/0022-3514.78.5.837>
- Helman, E., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017). The Unique Contributions of Perceiver and Target Characteristics in Person Perception. *Journal of Personality and Social Psychology*, 113(4), 513–529.
<https://doi.org/10.1037/pspa0000090>
- Henss, R. (1991). Perceiving age and attractiveness in facial photographs. *Journal of Applied Social Psychology*, 21(11), 933–946.
- Hönekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2), 199–209.
<https://doi.org/10.1037/0096-1523.32.2.199>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Jasienska, G., Lipson, S. F., Ellison, P. T., Thune, I., & Ziomkiewicz, A. (2006). Symmetrical women have higher potential fertility. *Evolution and Human Behavior*, 27(5), 390–400.
<https://doi.org/10.1016/j.evolhumbehav.2006.01.001>
- Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, 121(3), 313–323.
<https://doi.org/10.1016/j.cognition.2011.08.001>
- Jones, D., & Hill, K. (1993). Criteria of facial attractiveness in five populations. *Human Nature*, 4(3), 271–296. Retrieved from
<http://link.springer.com/article/10.1007/BF02692202>
- Kline, P. (1994). *An easy guide to factor analysis*. *Personality and Individual Differences*

(Vol. 17). London; New York: Routledge. [https://doi.org/10.1016/0191-8869\(94\)90040-X](https://doi.org/10.1016/0191-8869(94)90040-X)

- Lassek, W. D., & Gaulin, S. J. C. (2008). Waist-hip ratio and cognitive ability: Is gluteofemoral fat a privileged store of neurodevelopmental resources? *Evolution and Human Behavior*, 29(1), 26–34. <https://doi.org/10.1016/j.evolhumbehav.2007.07.005>
- Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social Psychology*, 82(6), 947–955. <https://doi.org/10.1037//0022-3514.82.6.947>
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, 51, 93–120. <https://doi.org/10.1146/annurev.psych.51.1.93>
- Maestriperi, D., Henry, A., & Nickels, N. (2017). Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behavioral and Brain Sciences*, 40. <https://doi.org/10.1017/S0140525X16000340>
- Marlowe, C. M., Schneider, S. L., & Nelson, C. E. (1996). Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased? *Journal of Applied Psychology*, 81(1), 11–21. <https://doi.org/10.1037/0021-9010.81.1.11>
- Matts, P. J., Fink, B., Grammer, K., & Burquest, M. (2007). Color homogeneity and visual perception of age, health, and attractiveness of female facial skin. *Journal of the American Academy of Dermatology*, 57(6), 977–984. <https://doi.org/10.1016/j.jaad.2007.07.040>
- Møller, A. P. (1999). Developmental stability is related to fitness. *The American Naturalist*, 153(5), 556–560. <https://doi.org/10.1086/303197>
- Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory*. McGraw-Hill Higher Education

(Vol. 3). <https://doi.org/10.1037/018882>

- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods*, 32(3), 396–402. <https://doi.org/10.3758/BF03200807>
- O'Toole, A. J., & Natu, V. (2013). Computational perspectives on the other-race effect. *Visual Cognition*, 21(9–10), 1121–1138. <https://doi.org/10.1080/13506285.2013.803505>
- Oldmeadow, J. A., Sutherland, C. A. M., & Young, A. W. (2013). Facial stereotype visualization through image averaging. *Social Psychological and Personality Science*, 4(5), 615–623. <https://doi.org/10.1177/1948550612469820>
- Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, 18(11), 566–70. <https://doi.org/10.1016/j.tics.2014.09.007>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32), 11087–92. <https://doi.org/10.1073/pnas.0805664105>
- Perrett, D. I., Penton-Voak, I. S., Little, A. C., Tiddeman, B. P., Burt, D. M., Schmidt, N., ... Barrett, L. (2002). Facial attractiveness judgements reflect learning of parental age characteristics. *Proceedings. Biological Sciences / The Royal Society*, 269(1494), 873–880. <https://doi.org/10.1098/rspb.2002.1971>
- Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., & Macrae, C. N. (2011). Stereotype-based modulation of person perception. *NeuroImage*, 57(2), 549–557. <https://doi.org/10.1016/j.neuroimage.2011.05.004>
- Quinn, K. a, & Macrae, C. N. (2005). Categorizing others: The dynamics of person construal. *Journal of Personality and Social Psychology*, 88(3), 467–79. <https://doi.org/10.1037/0022-3514.88.3.467>

- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, 57, 199–226. <https://doi.org/10.1146/annurev.psych.57.102904.190208>
- Santos, I. M., & Young, A. W. (2005). Exploring the perception of social characteristics in faces using the isolation effect. *Visual Cognition*, 12(1), 213–247. <https://doi.org/10.1080/13506280444000102>
- Santos, I. M., & Young, A. W. (2008). Effects of inversion and negation on social inferences from faces. *Perception*, 37(7), 1061–1078. <https://doi.org/10.1068/p5278>
- Santos, I. M., & Young, A. W. (2011). Inferring social attributes from different face regions: Evidence for holistic processing. *Quarterly Journal of Experimental Psychology*, 64(4), 751–66. <https://doi.org/10.1080/17470218.2010.519779>
- Simpson, J. A., & Gangestad, S. W. (1992). Sociosexuality and romantic partner choice. *Journal of Personality*, 60(1), 31–51. <https://doi.org/10.1111/j.1467-6494.1992.tb00264.x>
- South Palomares, J. K., Sutherland, C. A. M., & Young, A. W. (2018). Facial first impressions and partner preference models: Comparable or distinct underlying structures? *British Journal of Psychology*, 109, 538–563. <https://doi.org/10.1111/bjop.12286>
- South Palomares, J. K., & Young, A. W. (2017). Facial first impressions of partner preference traits: trustworthiness, status, and attractiveness. *Social Psychological and Personality Science*. <https://doi.org/10.1177/1948550617732388>
- Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, 127(1), 105–18. <https://doi.org/10.1016/j.cognition.2012.12.001>
- Sutherland, C. A. M., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and

- facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. <https://doi.org/10.1016/j.cognition.2016.09.006>
- Sutherland, C. A. M., Rhodes, G., & Young, A. W. (2017). Facial image manipulation: A tool for investigating social perception. *Social Psychological and Personality Science*, *8*(5), 538–551. <https://doi.org/10.1177/1948550617697176>
- Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, *108*(2), 397–415. <https://doi.org/10.1111/bjop.12206>
- Symons, D. (1979). *The evolution of human sexuality* (1st ed.). Oxford University Press.
- Tatarunaite, E., Playle, R., Hood, K., Shaw, W., & Richmond, S. (2005). Facial attractiveness: A longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*, *127*(6), 676–682. <https://doi.org/10.1016/j.ajodo.2004.01.029>
- Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(4), 42–50. <https://doi.org/10.1109/38.946630>
- Todd, P. M., Penke, L., Fasolo, B., & Lenton, A. P. (2007). Different cognitive processes underlie human mate choices and mate preferences. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(38), 15011–15016. <https://doi.org/10.1073/pnas.0705290104>
- Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>

- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*(6), 813–833.
<https://doi.org/10.1521/soco.2009.27.6.813>
- Todorov, A., & Porter, J. M. (2014). Misleading first impressions: Different for different facial images of the same person. *Psychological Science, 25*(7), 1404–1417.
<https://doi.org/10.1177/0956797614532474>
- van Rijsbergen, N., Jaworska, K., Rousselet, G. A., & Schyns, P. G. (2014). With age comes representational wisdom in social signals. *Current Biology, 24*(23), 2792–2796.
<https://doi.org/10.1016/j.cub.2014.09.075>
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*(3), 321–327. <https://doi.org/10.1007/BF02293557>
- Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences of the United States of America, 111*(32), E3353–61.
<https://doi.org/10.1073/pnas.1409860111>
- Walster, E., Aronson, V., Abrahams, D., & Rottman, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology, 4*(5), 508–516. <https://doi.org/10.1037/h0021188>
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592–598.
<https://doi.org/10.1111/j.1467-9280.2006.01750.x>
- Young, A. W. (2018). Faces, people and the brain: The 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology, 71*(3), 569–594.
<https://doi.org/10.1177/1747021817740275>
- Zentner, M. R., & Mitura, K. (2012). Stepping out of the caveman's shadow: Nations' gender

gap predicts degree of sex differentiation in mate preferences. *Psychological Science*, 23(10), 1176–85. <https://doi.org/10.1177/0956797612441004>