

SCIENTIFIC REPORTS



OPEN

Subjectivity and complexity of facial attractiveness

Miguel Ibáñez-Berganza¹, Ambra Amico¹ & Vittorio Loreto^{1,2,3}

The origin and meaning of facial beauty represent a longstanding puzzle. Despite the profuse literature devoted to facial attractiveness, its very nature, its determinants and the nature of inter-person differences remain controversial issues. Here we tackle such questions proposing a novel experimental approach in which human subjects, instead of rating natural faces, are allowed to efficiently explore the face-space and “sculpt” their favorite variation of a reference facial image. The results reveal that different subjects prefer distinguishable regions of the face-space, highlighting the essential subjectivity of the phenomenon. The different sculpted facial vectors exhibit strong correlations among pairs of facial distances, characterising the underlying universality and complexity of the cognitive processes, and the relative relevance and robustness of the different facial distances.

The notions of body beauty and harmony of proportions have fascinated scholars for centuries. From the ancient Greek canons, a countless number of studies have focused on unfolding what is behind the beauty of the face and the body. Nowadays the notion of facial beauty is a fast expanding field in many different disciplines including developmental psychology, evolutionary biology, sociology, cognitive science and neuroscience^{1–5}. Still, despite a profuse and multi-disciplinary literature, questions like the very nature of facial attractiveness, its determinants, and the origin of inter-subject variability of aesthetic criteria, elude a satisfactory understanding. Here, we revisit the question drawing conclusions based on an empirical approach through which we allow human subjects to “sculpt” their favorite facial variations by navigating the so called face-space and converging on specific *attractors*, or preferred regions in the face-space.

The face is the part of the human body from which we infer the most information about others, such as: gender, identity, intentions, emotions, attractiveness, age, or ethnicity^{6–8}. In particular, looking at a face, we are able to immediately acquire a consistent impression of its attractiveness. Still, we could have a hard time explaining what makes a face attractive to us. As a matter of fact, which variables determine attractiveness and their interactions are still poorly understood issues³.

Many works have been devoted to assessing the validity of the *natural selection hypothesis*, or beauty as a “certificate” of good phenotypic condition⁷. According to this hypothesis, a face is judged on average as attractive according to a set of innate rules typical of the human species, which stand out with respect to other social or individual factors. Some degree of consensus has, indeed, been reported^{9–13}. Most of these experiments are based on the measurement of correlations among *numerical ratings* assigned to a set of natural (or synthetic^{14,15}) facial images by raters belonging to different cultural groups. Much work in this field has also been devoted to assessing the covariation of the perceived beauty of a face with facial *traits* that are believed to signal good phenotypic condition, mainly: facial symmetry, *averageness* and secondary sexual traits. After decades of intense research, the role played by these traits is known to be limited: facial beauty seems to be more complex than symmetry⁵, *averageness*^{14,16} and secondary sexual traits^{7,17}.

Indeed, it has been documented that cultural, between-person and intra-person differences influence attractiveness perception in various ways⁴. As a representative example, the link between masculinity and attractiveness in male faces is subject to significant inter- and intra-subject differences^{4,5,7,18}. An evolutionary explanation is that exaggerated masculinity could be perceived as denoting a lack of some personality facets such as honesty or expressiveness¹⁵. In this context, the so called *multiple fitness* or *multiple motive model*^{4,11,19} proposes that attractiveness varies according to a variety of *motives*, each one evoking a different abstract attribute of the person whose face is evaluated.

On the other hand, an impressive amount of work is committed to the automatic facial beauty rating. This is tackled as a supervised inference problem whose training database is composed of natural facial images codified

¹Sapienza University of Rome, Physics Department, Piazzale Aldo Moro 2, 00185, Rome, Italy. ²Sony Computer Science Laboratories, Paris, 6, rue Amyot, 75005, Paris, France. ³Complexity Science Hub, Josefstadt Strasse 39, A 1080, Vienna, Austria. Correspondence and requests for materials should be addressed to M.I.-B. (email: miguel.berganza@roma1.infn.it)

Received: 24 January 2019
Accepted: 13 May 2019
Published online: 10 June 2019

by *vectors of facial coordinates in face-space*^{3,20,21}, along with (inter-subject averaged) numerical ratings assigned to them by human subjects, to be inferred. Works differ mainly on the codification of faces in the face-space: from a *geometric* face description (2D or 3D spatial coordinates of the *facial landmarks*), to a detailed description of the *texture* or luminosity degrees of freedom that provide a cue to the facial shape in depth (there also exist *holistic* representations, extracting lower-dimensional, non-local information from the facial image according to some criterion (Principal Component *eigenfaces* or Gabor filters); or using richer techniques which integrate geometric from skin textural and reflectivity characteristics). With the advent of deep hierarchical neural networks, the raw facial image is given as an input to the algorithm, which automatically extracts the putative relevant features in the inference process, although in a hardly accessible way (the *black box problem*).

The supervised inference of ratings may help to address, albeit indirectly, the impact of various facial features on attractiveness. Although the relative relevance of different features has been discussed in various articles, robust conclusions are lacking^{3,22–28}. The results about the relative relevance of the *kind* (geometric, textural and holistic) of facial attributes to attractiveness are controversial as well^{3,29–33}. In any case, the integration of different kinds of variables seems to improve the inference results^{29,34}, suggesting that these are complementarily taken into account in the cognitive process of attractiveness assessment.

Facial beauty is, hence, probably not a universal function of a set of few facial properties, as implicitly assumed in many references, but the result of a complex process in which multiple semantic concepts, providing cues to personality facets, are inferred. The literature concerning inference of personality traits indicates that such semantic concepts may be encoded in global combinations of facial features, in a complex way³⁵. This motivates a study of facial beauty beyond the subject-averaged rating, focusing on the inter-subject heterogeneity and on the global combinations of various facial features generating such a diversity.

In summary, the complexity of facial attractiveness perception so far prevented a satisfactory understanding of how attractiveness relates to various facial elements³, and of the nature of inter-personal differences. In order to make progress, from a methodological point of view it is important to highlight three key factors. (A) The possible mutual influence among geometric, texture and detailed features³⁶. Even considering the problem in terms of geometric variables only, the possible existence of *interactions* or mutual dependencies between different facial components may induce a variety of possible pleasant faces, even for the single subject. (B) The undersampling of the relevant face-space, due to the many different prototypes of facial beauty^{14,29}. (C) The subjectivity of the phenomenon, probably hindered by the use of the average numerical beauty ratings. The complexity and richness of the perceptual process, suggested by the multiple-motive hypothesis and by previous work about perception of personality dimensions^{6,37–39}, eludes a description in terms of average ratings, a quantity that has already been observed to be inadequate³.

In light of these considerations, we here address the phenomenon of facial preference through an empirical approach that aims at removing the biases of ratings, focusing instead on the possibility given to human subjects to freely explore a suitably defined face-space. By means of a dedicated software, based on image deformation and genetic algorithms, we focus on inter-subject differences in aesthetic criterion and let several subjects sculpt their favorite variation of a *reference portrait*, parametrized by a vector of geometric facial coordinates. We observe how different subjects tend to systematically sculpt facial vectors in different regions of the face-space, which we call *attractors*, pointing towards a strong subjectivity in the perception of facial beauty. In addition, the facial vectors sculpted by different subjects exhibit strong correlations for pairs of facial distances, which is a manifestation of the underlying universality and complexity of the cognitive process of facial image discrimination. The correlations contain information regarding the different sources of variability in the dataset of selected vectors. For instance, though a difference between male-female subjects is clearly observed, the largest differences among facial variations, elicited by a principal component analysis, result from criteria that are transversal with respect to the gender only. A third important result concerns the assessment of the robustness of the results with respect to the degrees of freedom not described in the face-space. Crucially, in our approach, the luminance, texture and detailed degrees of freedom are decoupled from the geometric features defining the face-space, and deliberately kept *fixed, and common for all the subjects*. Finally, we observe that the overall experimental results are, interestingly, partially robust and independent of the detailed degrees of freedom (the reference portrait).

The current experimental scheme bypasses the three confounding factors (A–C) mentioned in the precedent paragraph. (A) Uncontrolled sources of biases are absent in our study, since all possible facial variations (given the reference portrait) are described by points in the face-space. (B) In our face-space of reduced dimensionality and unchanged texture degrees of freedom the undersampling is mitigated, making possible an efficient exploration of the face space and allowing for an accurate characterisation of the single-subject attractor. (C) This allows us to fully account for subjectivity: we are able to analyse the differences among different subject's preferred facial modifications.

Results

Preferred facial images as extrema in face-space. We consider a face-space defined by a set of geometric coordinates illustrated in Fig. 1A. A face is parametrized in terms of a set of 10 non-redundant Cartesian coordinates of 7 single landmarks $\vec{\ell}_\alpha = (x_\alpha, y_\alpha)$ or, alternatively, in terms of a vector of $D = 11$ inter-landmark distances $\mathbf{d} = (d_i)_{i=1}^D$. The face-space vector components f_i are, in this way, either landmark Cartesian coordinates or inter-landmark distances. From a vector of facial coordinates \mathbf{f} and a *reference facial portrait* corresponding to a real person, we then construct a facial image by a continuous deformation of the reference portrait such that its landmark geometric coordinates acquire the desired value, \mathbf{f} (Fig. 1B,C). Within a single experiment, the reference portrait (the image texture) is unchanged and only the geometric position of the landmarks can change (for an in-depth explanation see Sec. Methods and the Supplementary Information).

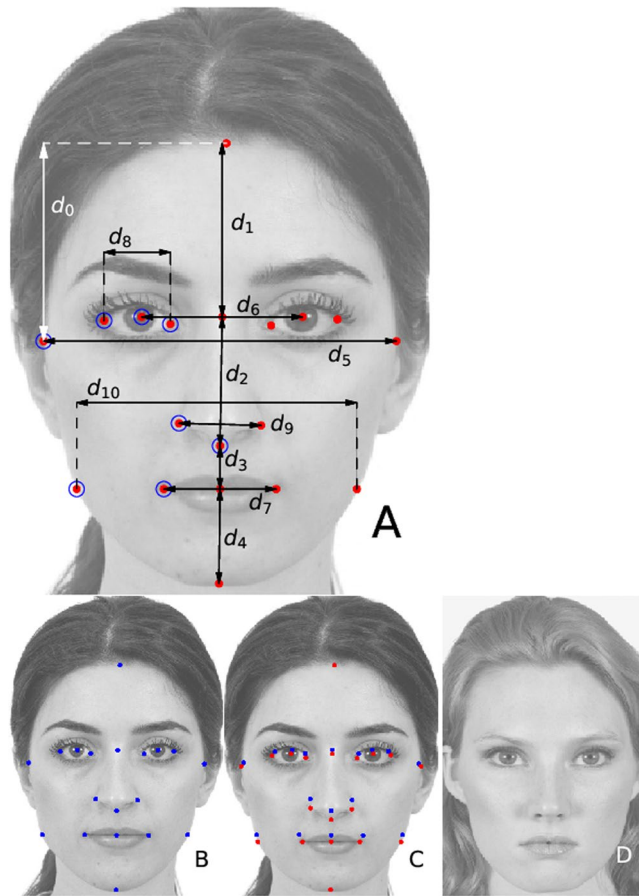


Figure 1. (A) The parameters defining the face space. The red points indicate the *landmarks*, $\alpha = 1, \dots, 18$, whose 2D varying Cartesian coordinates generate the continuum of face space. The face space points are parametrised in terms of vectors \mathbf{f} whose components are the Cartesian coordinates of a set of non-redundant landmarks $\bar{\mathcal{L}}_\alpha$ (signaled with an empty circle), or in terms of (vertical or horizontal) distances d_i ($i = 0, \dots, 10$) among some pairs of landmarks $d_i = |\mathbf{x}_{\alpha(i)} - \mathbf{x}_{\beta(i)}|$ or $d_i = |\mathbf{y}_{\alpha(i)} - \mathbf{y}_{\beta(i)}|$ (arrows). (B) Reference portrait RP1 used in experiment E1 along with its corresponding landmarks (in blue). (C) Image deformation of RP1 according to a given vector of inter-landmark distances \mathbf{d} : the blue reference portrait landmarks are shifted (leading to the red points) so that their inter-landmark distances are \mathbf{d} , and the reference image (B) is consequently deformed. (D) Image deformation of the reference portrait RP2 according to the same vector of distances \mathbf{d} as in (C).

The aim of the experimental method is to provide a *population* of N facial vectors, $\{\mathbf{f}^{(s,n)}\}_m$, with $n = 1, \dots, N$ and $\mathbf{f}^{(s,n)} \in \mathbb{R}^D$, for each experimental subject, s . Such a population is considered as an empirical sample of the subject's attractor, or the face-space region of his/her preferred modifications of the reference portrait. This means that the subject would probabilistically prefer facial images associated with vectors that are close to the attractor, rather than local fluctuations away from it (for a precise definition see the Supplementary Section S2). In our experimental scheme, the subject does not sculpt the population by successive discrimination among faces differing by a single coordinate, which turns out to be an inefficient strategy of face-space exploration, but rather through the interaction with a genetic algorithm (see sections Methods, Supplementary Section S3).

In a first experiment (E1), we have let $S_1 = 95$ subjects sculpt their facial variations of reference portrait RP1 (see 1-A). This results in a final population, $\mathcal{S}_1 = \{\mathbf{f}^{(s,n)}\}_{s=1, n=1}^{S_1, N}$ of $N = 28$ facial vectors for each subject. Starting from N initial random facial vectors, the FACEXPLORE software generates pairs of facial images that are presented to the subject, who selects the one that he/she prefers. Based on N left/right choices, a genetic algorithm produces a successive generation of N vectors, in a constant feedback loop of offspring generation and *selection* operated by the subject. The iteration of this process leads to a sequence of T generations of facial vectors, each one more adapted than the last to the subject's selection criteria, eventually converging to a pseudo-stationary regime in which the populations are similar to themselves and among consecutive generations. Figure 2 reports the evolution (versus the generation index, $t = 1, \dots, T = 10$) of the *intra-population distance*, the distance among faces within the single populations sculpted by 10 different, randomly chosen, subjects in E1 (see Supplementary Section S4 for details). In the next subsection, we discuss the degree of reproducibility of our results as a function of N , T and S_1 .

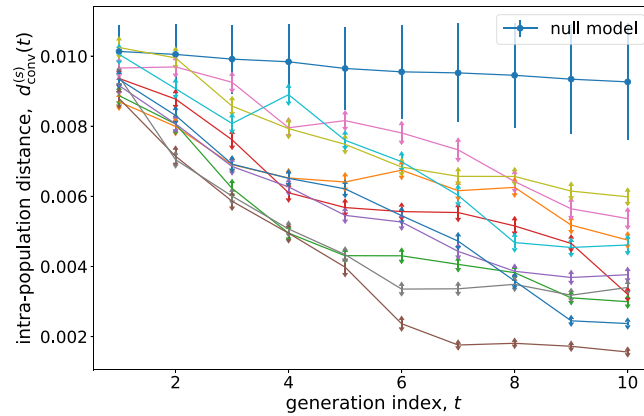


Figure 2. Intra-population distance of the populations sculpted by different subjects (s) as a function of the generation (t). The Euclidean metrics in face space has been used (see Supplementary Sec. S4), although the results are qualitatively equal for other relevant metrics. Each curve corresponds to a different subject (for 10 randomly chosen subjects). The upper curve of joined circles corresponds to the null model genetic experiment, in which the left/right choices are random.

The intra-population distance decreases with the generation index, indicating that the populations sculpted by single subjects tend to clusterize in a region of the face-space. This clustering is not observed in a null experiment in which the left-right decisions are taken randomly. Remarkably, a diversity of behaviors towards the pseudo-stationary regime is observed, already signaling differences in the way the face-space is explored.

From now on, we will consider the final population sculpted by the s -th subject, $\{\mathbf{f}^{(s,n)}\}_{n=1}^N$, as the final, $T=10$ -th generation of the sequence of populations sculpted by this subject in E1. In the next subsection we show that the face-space attractors of different subjects are actually significantly and consistently different. This experimental scheme is, therefore, able to *resolve* the subjective character of attractiveness, as the single subject tends to sculpt populations of vectors clustered in a narrow region in the face-space in successive realisations of the experiment. All these facts imply that the single subject attractor can be operationally characterised as an *extremum of a subject-dependent, probabilistic function in face-space*, which may be inferred from the populations sculpted by the subject in several instances of the experiment (see Supplementary Section S2 for a complete definition). The attractors are extrema of such a function in the sense that a significant fluctuation of a vector coordinate away from its value in the attractor will tend to lower its probability of being selected by the subject, given the reference portrait.

Assessment of subjectivity: distinguishable aesthetic ideals. In order to assess the subjectivity of the sculpting process, we need to measure to what extent the same subject, by repeating the same experiment, would sculpt populations of facial vectors closer to each other than to populations sculpted by distinct subjects. To this end we performed a second experiment (E2), in which a subset of $S_{sc}=6$ subjects were asked to perform $m=6$ instances of an experiment E1, with the common reference portrait RP1, different (random) initial conditions and sequence of random numbers in the genetic algorithm. The subjectivity is assessed through the comparison of two sets of distances: (i) the $(S_{sc}m(m-1)/2)$ *self-consistency distances* among facial populations sculpted by the same subject in different instances of the experiment E2; (ii) the $(S_1(S_1-1)/2)$ *inter-subject distances* between couples of populations sculpted by different subjects in experiment E1 (see Supplementary Section S4 for details). If subjectivity was at play in the sculpting process, and not hindered by the stochasticity of the algorithm, the self-consistency distances would be lower than inter-subject distances.

This is clearly the case, see Fig. 3: self-consistency distances are lower than inter-subject distances (Student's $p < 10^{-30}$). In Fig. 3 we also report the histogram of *intra-population distances*, i.e., the average distance among the vectors belonging to a population, for different populations sculpted by different subjects in E1 (blue curve). The intra-population distances are not suitable for an assessment of the subject self-consistency, since they strongly depend on the number of generations performed by the genetic algorithm (c.f. Fig. 2). The emerging scenario is that of single subjects who, in a single realization of the sculpting experiment, end up in a very clustered population (blue curve in Fig. 3). Performing several realizations of the same experiment leads the subject to a slightly different population in face-space (orange curve in Fig. 3, labelled “self-consistency”). These self-consistent populations are anyway closer to each other than to populations sculpted by different subjects, as witnessed by the larger *inter-subject distances*, whose histogram is presented in the green curve in Fig. 3. A crucial point is that the distance between the inter-subject (green curve, i) and self-consistency (orange curve, sc) histograms in Fig. 3, $t = (\mu_i - \mu_{sc}) / (\sigma_i^2 + \sigma_{sc}^2)^{1/2} = 0.82(1)$ (see Supplementary Fig. S3) would be *even larger* in an experiment with a higher number of generations T . Using larger values of the genetic algorithm parameters T and N would result in a lower value of μ_{sc} , at the cost of a larger experimental time, since NT binary choices are required from the subject (see Sec. Methods and Supplementary Sec. S3). Furthermore, larger values of S_1 , S_{sc} , m would give rise to a lower statistical error of the considered observables (see Supplementary Sec. S4), proportional to $1/\sqrt{S_{1,sc}}$ and, in particular, to an even more significant difference among both histograms, since the uncertainty of their average is

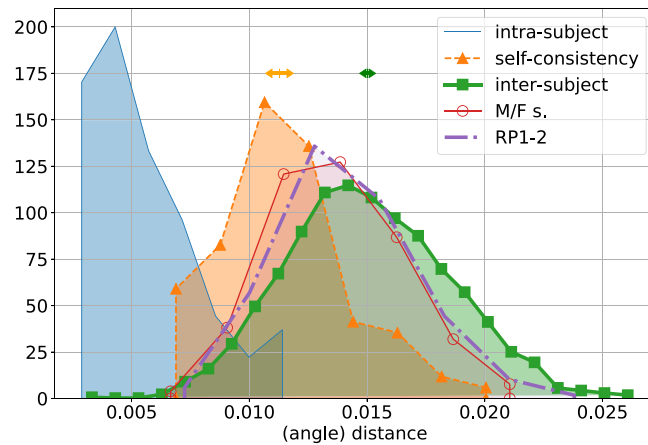


Figure 3. Main panel: Normalised histograms of pseudo-distances. **Blue:** subject intra-population distances, or self-distances of all the populations sculpted in E1. **Orange:** self-consistency distances, or distances among couples of populations sculpted by the same subject in E2. **Green:** inter-subject distances, or distances among couples of populations sculpted by different subjects in E1. **Purple:** distances among couples of populations sculpted by different subjects in different experiments, E1 and E3 (differing in the reference portrait). **Red:** distances among couples of populations sculpted by subjects of different gender in E1. The orange and green arrowed segments over the self-consistency and inter-subject histograms indicate the confidence intervals of the histogram averages, $\mu_{sc} \pm \sigma_{sc}/n_{sc}^{1/2}$ and $\mu_i \pm \sigma_i/n_i^{1/2}$ respectively, with $n_{sc} = S_{sc}m(m-1)/2$ and $n_i = S_i(S_i-1)/2$.

proportional to σ_i/S_i and $\sigma_{sc}/\sqrt{S_{sc}m}$, respectively. In any case, the values used in experiments E1-2 are large enough to assess the differences among different subjects' attractors in a significant way.

The set of populations $S_1 = \{\mathbf{f}^{(s,n)}\}_{s,n}$ sculpted in E1 exhibits facial coordinates which vary in a wide range: roughly 0.018(10) per coordinate of the total face length, corresponding to ~ 3.2 mm in the average female face⁴⁰ (see the average $\langle \mathbf{f} \rangle$ and standard deviation σ of the single coordinates in Supplementary Fig. S5). The self-consistency distance $\mu_{sc} \pm \sigma_{sc}$, with which the experiment allows to resolve the single-individual attractor is, remarkably, much lower, equal to 0.0067(18) per coordinate (using the simple Euclidean-metrics in face-space, see Supplementary Section S8), barely twice the pixel image resolution, $\sim 400^{-1}$ (Figure Supplementary Section S4). This quantity corresponds to 1.18(30) mm in the female average facial length.

Several metrics among facial vectors have been used to compute the inter-subject and self-consistency distances: *Euclidean*, *Mahalanobis*, *angle*- and *Byatt-Rhodes metrics* (see Supplementary Section S4 and^{20,21}). The angle-metric (the angle subtended among standardised Principal Components (PC's) in face-space) turns out to be the one with which the statistical distinction is more significant (see Supplementary Fig. S3, and subsection "Differences induced by the subject gender" for the definition of PC's). This result is compatible with previous work proposing that such face-space metrics is the one that best captures differences in facial identity^{21,41}. Further results regarding the *t*-value difference among both histograms as a function of the face-space metrics can be found in the Supplementary Section S4. Using the simple Euclidean metrics (the Euclidean distance per coordinate in physical coordinates), the inter-subject and self-consistency distances result slightly more overlapping, although still clearly distinct. For the sake of the statistical discernibility among the inter-subject and self-consistency distances, it is observed that the 10 dimensions involved in the definition of the face space are redundant in the sense that defining the face-space metrics in terms of the 7 most varying PC's, the two sets of distances result more significantly different (see Supplementary Fig. S3). For completeness, in Fig. 3 we also report two further sets of distances. The red line histogram corresponds to pseudo-distances among pairs of populations sculpted by subjects of different gender in E1, while the purple line histogram corresponds to the pseudo-distances among pairs of populations sculpted by different subjects with different reference portraits (E3, see "Relevance of facial features", before).

These findings highlight the intrinsic subjectivity of facial attractiveness. Despite the limited freedom of choice, the reduced dimension of the face-space, and the common reference portrait, single subjects tend to sculpt a region of face-space that is systematically closer to their previous selections than to other subjects' sculptures. Indeed, the probability of two facial vectors sculpted by the same subject to be closer than two facial vectors sculpted by different subjects in E1 is $p_{12} = 0.79(1)$ (see Supplementary Section S8).

A further interesting observation about Fig. 3 concerns the overlap between the histograms of *self-consistency* and *inter-subject distances*. Its existence allows us to reconcile the strong subjectivity unveiled by experiments E1-2, and the universality reported in the literature. The couples of facial vectors which are involved in distances for which there is a high overlap correspond to commonly preferred faces, around the most probable vector in the dataset, $\langle \mathbf{f} \rangle$. Within a low experimental precision, or an accuracy larger than the standard deviation per coordinate $a > |\sigma|/D$, all the subjects appear to agree in their choices. Under this perspective, the reported universality of beauty could be the side-effect of an experimental procedure where subjects express their preferences among a limited set of predefined options, the real facial images, in a high-dimensional face-space (indeed, the effective number of relevant facial dimensions may be of the order of hundreds⁴²). In such an undersampling

situation, different natural faces exhibit very different number of facial coordinates g_i (or, more precisely, of PC's, see before), close to the most probable value $\langle g_i \rangle$, with respect to their standard deviation (say, $\sigma(g_i)$). The faces exhibiting many coordinates in the commonly preferred region are consensually preferred, and most highly rated²⁰. By letting the subjects sculpt instead their preferred modification in a lower-dimensional face space, as in experiments E1-2, the subjects exclude extreme values of the coordinates, and manage to fine-tune them according to their personal criterion. In this circumstance, it is possible to resolve the subjects' preferences with higher accuracy, $\mu_{sc} < |\sigma|/D$, unveiling a strong subjectiveness. Our data suggest that the higher the accuracy with which the single subject attractor is resolved, the more distinguishable different subjects' attractors result in the face space. This picture suggests a *complete subjectivity*, or complete distinctiveness of different subjects' criteria (see also Sec. Methods).

Correlations among different facial features. In our experimental scheme, only geometric degrees of freedom may change. This allows us to determine the personal attractors efficiently and accurately, in a not too high-dimensional face-space. Moreover, it avoids the uncontrolled influence of features not described in the face-space. However, as anticipated in Sec. Introduction, it is also essential in this framework to account for possible mutual dependencies between different components of the facial vectors.

Besides the average and standard deviation of single coordinates referenced above, a quantity of crucial importance, despite the scarce interest that the literature has dedicated to it, is the correlation among facial coordinates from subject to subject. We denote with \mathbf{y} the standardised fluctuations of the vector \mathbf{f} around the experimental average, $y_i = (f_i - \langle f_i \rangle)/\sigma_i$. The sculpted facial vectors presenting a fluctuation of a coordinate y_i (say, a larger mouth width, $y_7 > 0$ in terms of inter-landmark distances) *tend to consequently present positive and negative fluctuations of other facial coordinates* $y_{j \neq i}$ (e.g., a higher mouth, $y_4 > 0$). The sign and magnitude of such covariations is given by the *correlation matrix* among fluctuations of facial coordinates. This is the positive definite, symmetric matrix $C_{ij} = \langle y_i y_j \rangle$, averaged over subjects $\langle \cdot \rangle = \sum_{s=1}^S \cdot / S$. In order to subtract the influence of correlations within the single-subject attractor, only one population vector, of index $n_b(s)$, uncorrelated and randomly distributed, is considered for each subject s ; the average and standard deviation of the matrix elements C_{ij} have been obtained from many bootstrapping realisations, labelled by b , of the indices $n_b(s)$, see Supplementary Section S4. The experimental matrix C exhibits a proliferation of non-zero elements (32% of the matrix elements presenting a p -value $< 5 \cdot 10^{-2}$, see Supplementary Section S11), unveiling the presence of strong correlations among several couples of facial coordinates.

The most strongly correlated C elements are among vertical or horizontal distances (see Supplementary Fig. S9 and Table S4). Such strong correlations are easily interpretable: wider faces in \mathcal{S}_1 tend to exhibit larger inter-eye distances and wider mouths and jaws; higher nose endpoints, in their turn, covary with higher mouths and eyes; higher eyes covary with higher mouths, and so on. Perhaps the most remarkable aspect of the matrix C is the proliferation of couples of vertical-horizontal coordinates, highlighting the crucial role played by *oblique* correlations. The sign of oblique correlations C_{ij} (see Supplementary Table S4) is such that fluctuations of a landmark position \vec{l}_α covary with fluctuations of different landmarks \vec{l}_β in such a way some inter α, β -landmark segment slopes are restored with respect to their average value. This is so for the most correlated couples of vertical-horizontal coordinates i, j ($p < 5 \cdot 10^{-2}$).

The information brought by the correlation matrix helps in this way to construct a remarkably clear picture of the experimental distribution of facial vectors. The inter-subject differences and the experimental stochasticity induce fluctuations around the average facial vector $\mathbf{y} = \mathbf{0}$. The fluctuations are, however, strongly correlated in the facial coordinates, in such a way that vertical and horizontal coordinates covary positively and, *at the same time*, the value of some inter-landmark segment slopes shown in Fig. 4, of prominent relative importance, do not change too much with respect to their average value (see Supplementary Section S13).

These findings indicate that, for a meaningful inference of the perceived attractiveness in face-space, one should consider the impact of at least *linear combinations* of facial coordinates, rather than the impact of single facial coordinates. The intrinsic complexity of attractiveness perception cannot be satisfactorily inferred through a simple regression of facial datasets using a sum of functions of single facial coordinates (see also Supplementary Section S14 and⁴³).

Relevance of facial features: the variable hierarchy. In this section we discuss the robustness of the results presented above. One of the crucial questions in facial attractiveness is what is the relevant set of variables which mainly determine the perceived attractiveness of a face^{3,36}. A formulation of the problem in theoretical-information terms is that of finding a hierarchy of relevant facial features. It is such that, when enriching the description with more variables in high levels of the hierarchy, the resulting variables in lower levels result unchanged. In the present study, the geometric quantities can be considered as low-level variables in the extent to which they are not influenced by the reference portrait, or by the luminance and texture facial features that have been disregarded and kept unchanged in the face-space description.

To settle this question, we performed a third experiment, dubbed E3, in which we asked the \mathcal{S}_1 participants in E1 to repeat the experiment using a different reference portrait (RP2, see Fig. 1D). Afterwards, we have compared the resulting set of sculpted facial vectors, \mathcal{S}_3 , with the outcome of experiment E1, \mathcal{S}_1 . Interestingly, a statistical t -test shows that, while some facial coordinates result clearly distinguishable, others result statistically indistinguishable, signifying their robustness with respect to the texture facial features determined by the reference portrait. These are, in terms of inter-landmark distances, d_p , the coordinates $d_{2,6,7,10}$, indistinguishable with $p > 0.1$ (see Supplementary Fig. S6). If, instead of focusing on the distribution of single quantities y_p , one considers instead the correlations, $y_i y_j$, the results (see Supplementary Table S4) turn out to be robust within their statistical

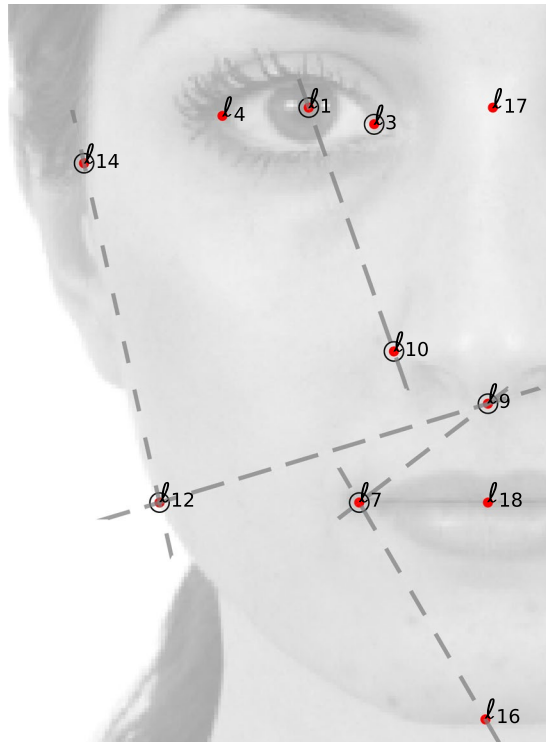


Figure 4. Relevant inter-landmark segments. The correlation matrix elements C_{ij} involving vertical and horizontal landmark coordinates, $\langle x_{\alpha(i)} y_{\alpha(j)} \rangle$ can be understood geometrically as a statistical invariance of the value of some inter-landmark segment slopes (dashed lines) with respect to their average value (represented in the figure). The sign of oblique C_{ij} 's coincide with that of the slope of the inter-landmark lines $\langle (y_{\alpha(i)} - \langle y_{\alpha(j)} \rangle) / (x_{\alpha(i)} - \langle x_{\alpha(j)} \rangle) \rangle$. For instance, the most correlated horizontal-vertical landmarks are $\langle x_{12} y_9 \rangle$, exhibiting a positive sign (c.f. Supplementary Table S4): indeed, for lower nose endpoints (which correspond to a positive fluctuation $y_9 > \langle y_9 \rangle$), the 9–12 angle can be restored only by increasing the x_{12} -coordinate, $x_{12} > \langle x_{12} \rangle$.

errors, since only 2% of the matrix elements C_{ij} result significantly distinguishable ($p < 0.075$, and none of them for $p < 0.05$).

The ensemble of these results implies a strong robustness of the results presented above, namely the subjectivity and the correlations among different facial features, with respect to a change in the reference portrait. It is remarkable that the coordinates $i = 2, 6, 7, 10$ in \mathcal{S}_1 are indistinguishable from those in \mathcal{S}_3 up to a remarkably small scale. For them, the average difference of couples of coordinates, $\langle f_i^{(s)} - f_i^{(s')} \rangle_{s,s'}$ (with subjects s, s' belonging to E1 and E3, respectively), vanishes up to small fluctuations, lower than the statistical error of such quantity. Such an error, of order $(S_1 S_3)^{-1/2}$, see Supplementary Section S10, is: $\sigma(f_i^{(s)} - f_i^{(s')}) = 1.54 \cdot 10^{-2}$ per coordinate, which corresponds to 0.27 mm in the average female face. We consider this result as one of the most remarkable of the present work. It highlights the striking robustness of the inter-landmark distances $d_{2,6,7,10}$. Such variables are, therefore, in low levels of the variable hierarchy, suggesting that they have prominent and intrinsic importance in the cognitive mechanism of face perception.

Differences induced by the subject gender. An extensively debated question in the literature is to what extent the subject gender influences attractiveness, a question that the present experimental scheme is particularly suited to address. Partitioning the dataset accordingly, $\mathcal{S}_1 = \mathcal{S}_m \cup \mathcal{S}_f$, it is obtained that, again, some facial coordinates are barely distinguishable or completely indistinguishable in both sets ($d_{3,4,6,7}$, see Supplementary Fig. S7). Conversely, some coordinates are noticeably distinguishable. Compared to female subjects, male subjects tend to prefer thinner faces and jaws ($d_{5,10}$), lower eyes (d_1), higher zygomatic bones (d_0), larger eye width (d_8). The difference becomes very distinguishable along $d_{2,9}$ ($p < 3 \cdot 10^{-3}$, Supplementary Fig. S7): males definitely prefer shorter and thinner noses. These results are partially in agreement with previous findings in the literature, that highlight male subjects' preference for smaller lower face area and higher cheekbones^{14,44}. Furthermore, they also provide accurate relative differences along each coordinate and reveal that, at least for the two reference portraits RP1-2, the facial feature leading to larger differences among men and women attractors is the nose.

A deeper insight is obtained by the analysis of PC's. These are the projections of the physical coordinates on the C-matrix eigenvectors, $\mathbf{y}' = \mathbf{E}\mathbf{y}$ (where $\mathbf{E}\mathbf{C}\mathbf{E}^\dagger = \text{diag}(\lambda_1, \dots, \lambda_D)$). The different principal components y'_i are, in other words, uncorrelated linear combinations of the physical coordinates $\langle y'_i y'_j \rangle = \lambda_i \delta_{ij}$. Principal components corresponding to large eigenvalues (as y'_{10}) represent the linear combinations of physical coordinates accounting for as much of the database variability, while those corresponding to the lowest eigenvalues represent

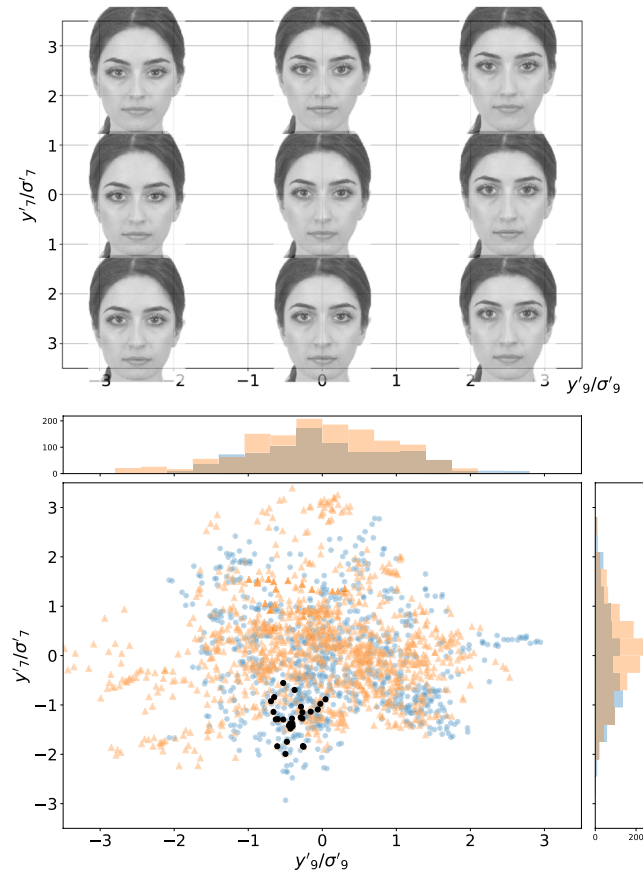


Figure 5. Top figure: facial images corresponding to the deformation of the average facial vector along two different principal axes (the $\mathbf{e}^{(7)}$, $\mathbf{e}^{(9)}$ eigenvectors of the correlation matrix C , corresponding to the fourth and second larger eigenvalues, λ_7 , λ_9). The axes represent the principal components along these axes, (y'_9, y'_7) in units of their standard deviations ($\lambda_i^{1/2}$). In other words, the image is generated from the facial vector $\mathbf{y} = E^\dagger(y'_7 \mathbf{e}^{(7)} + y'_9 \mathbf{e}^{(9)})$. Bottom figure: selected facial vectors. Each point is a projection of a selected facial vector in the principal axes corresponding to the Top figure, i.e., each point has coordinates $y'_7^{(s,n)}$, $y'_9^{(s,n)}$, for all s, n in the E1 dataset. Blue points correspond to male subjects, and orange triangles to female subjects (male subjects tend to sculpt vectors with $y'_9 < 0$, and vice-versa). The black points correspond to a population sculpted by a single, randomly selected, subject.

the most improbable, or “forbidden” linear combinations of fluctuations away from the average $\mathbf{y} = \mathbf{0}$ (see the Supplementary Information). Different principal axes ($\mathbf{e}^{(k)}$, the rows of matrix E) describe the different, independent sources of variability in the dataset, that could reflect the subjects’ traits most distinguishing their aesthetic criteria (as the gender).

It turns out that faces corresponding to different subject’s gender are distinguishable on three PC’s (see Supplementary Fig. S8). Quite interestingly, such principal axes are not the ones exhibiting the largest eigenvalue, suggesting that the largest differences among selected faces correspond to inter-subject *criteria* that are transversal with respect to the subject’s gender. Figure 5 shows some image deformations of the average face along two principal axes: $\mathbf{e}^{(9)}$, $\mathbf{e}^{(7)}$ (the 2nd and the 3rd most variant eigenvectors of C). The PC defined by $\mathbf{e}^{(9)}$ is male/female distinguishable (males preferring negative values of y'_9). Instead, the y'_7 coordinate is gender-indistinguishable, and it could correspond to a different subject’s quality, as the predilection for assertiveness, neoteny, or a different personality dimension, in the language of the *multiple motive hypothesis*^{4,11,19}.

Discussion

In this article, we have introduced an experimental behavioural method that allows human subjects to efficiently select their preferred modification of a reference portrait in the multi-dimensional face-space (and, in principle, in general spaces of images that can be parametrised with 2D landmark coordinates). The method allows to flexibly and accurately determine the face-space regions which are representative of a given subject’s criterion. It opens the path to a novel, data-driven approach to cognitive research in face perception, allowing scholars to: (1) quantitatively address the inter-subject differences in the *resulting sculpted shapes, beyond the rating*; (2) isolate the influence of a *secondary* set of variables (such as texture features) and *a posteriori* address their influence (something that cannot be directly done with databases of natural facial images); (3) analyse a resulting set of facial vectors without being limited or conditioned by the *a priori* correlations present in natural image databases.

The method (based on our software FACEEXPLORE, whose details are explained in the Supplementary Information) permits a highly accurate description of the single subject or subject category preferences in the face-space, thanks to the geometric/texture separation of facial degrees of freedom and to a genetic algorithm for efficient search in the face space. Using this technique, we have performed a set of experiments in which the single subjects preferred region in the face space have been determined with an unprecedented accuracy, below the millimeter per facial coordinate.

Such experiments allow us to draw the following conclusions. First of all, attractiveness turns out to be associated with the existence of subject-dependent specific regions in the face space that we dubbed attractors, highlighting the essential subjectivity of attractiveness. Despite the limited face-space dimension, and the homogeneity of the statistical universe (composed of subjects of the same cultural group), different subjects clearly tend to prefer different facial variations, suggesting that the subjectivity should be taken into account for a complete scientific picture of the phenomenon. Larger databases and more heterogeneous statistical universes would only make the essential subjectivity of attractiveness perception even more evident.

In light of these facts, the validity of the natural selection hypothesis (universality, impact of averageness, symmetry and sexually dimorphic traits) may be arguably a matter of the precision of the length scale and of the facial image resolution of the facial description. Within a sufficiently accurate description of the subjects' criterion in face-space, the phenomenon emerges in its whole complexity, showing that the preferred faces of different subjects are systematically different among themselves and, consequently, different from the average face. In their turn, these differences reflect personal features and circumstances that condition the subject's preferences, one of which is the subject's gender.

The second important conclusion we can draw concerns the patterns associated to different subjects' attractors. Different sculpted facial vectors exhibit strong correlations among pairs of facial distances, characterising the underlying universality and complexity of the cognitive processes, leading, in its turn, to the observed subjectivity⁴. Our study reveals, in particular, the crucial importance of correlations among vertical and horizontal coordinates, whose existence and relevance have been, to the best of our knowledge, only postulated^{22,24,35}. Different facial variations are strongly correlated, a fact that confirms the holistic way in which we perceive faces³⁶. Our results suggest to consider attractiveness not as a scalar quantity, rather as the outcome of a complex process in which various semantic motives are evaluated. These are probably encoded in pairwise and higher-order correlations among facial features, more than in the value of single facial coordinates³⁵.

A third result concerns the role of the subject's gender in the assessment of attractiveness. This is, indeed, an important source of diversity in our dataset. Nose length and width, eye height, face and jawbone width, zygomatic bone height, turn out to be the main facial traits distinguishing male and female observers. However, a principal component analysis suggests that the largest differences among selected facial variants correspond to principal axes that are independent of the subject's gender. Abstract personality dimensions have been observed to be consensually attributed to faces, and the impact of such qualities on various facial elements have been measured through principal component analysis^{6,37–39}. Such principal axes could be correlated with those of the present study. This would be a confirmation of the postulated connection between attractiveness and personality judgments^{1,6,45}. It would allow to elicit the different traits that are judged by the subjects in a bottom-up, data-driven fashion.

A further noticeable result is the assessment of the influence of the reference portrait in the distribution of sculpted facial vectors. Quite remarkably, the *a priori* dimensionality reduction implicit in our analysis (ignoring texture degrees of freedom), turns out a posteriori to be sufficient and justified (see Sec. Methods).

In summary, the novel experimental approach proposed in this article allowed us to unveil the essential subjectivity of attractiveness. The subjectivity emerges more evidently in the present scheme, since the reduction of the number of face space dimensions allows to avoid the undersampling occurring in experiments in which the subjects are asked to choose or rate natural faces.

We believe that the generality and reliability of the present approach could have a strong impact on future studies about beauty and pleasantness in different domains.

Possible completions of the present work are: an assessment of the robustness of principal components; an analysis of the intra-subject correlation matrix of facial coordinates; a variant of the analysis of correlations in an experiment with real facial images (whose landmarks could be automatically identified with deep learning techniques⁴⁶); an unsupervised inference analysis of the database (already being carried on in our group) within the framework of the Maximum Entropy method.

Methods

Face space. Our experimental design is based on the parametrisation of the face in a 10-dimensional face-space defined by $D = 11$ vertical and horizontal *inter-landmark distances*, $\mathbf{d} = (d_i)_{i=0}^D$ between standard facial landmarks (see Fig. 1A). The inter-landmark distances are subject to a constraint $h = \sum_{i=1}^4 d_i = 1$, reflecting the intrinsic scale invariance of the problem, in such a way that all distances d_i are in units of the total facial length (i.e., they represent *proportions with respect to the facial length*, rather than absolute distances). As vector of facial coordinates \mathbf{f} , we have considered both the 11 distances $f_i = d_i$ themselves or, alternatively, the non-redundant (and unconstrained) subset of $D = 10$ *Cartesian landmark coordinates* of a set of landmarks $\vec{\ell}_\alpha = (x_\alpha, y_\alpha)$ (with $\alpha = 1, 3, 7, 9, 10, 12, 14$, see Fig. 4 and Supplementary Sec. S13), that can be unambiguously retrieved from the set of inter-landmark distances. All the results presented in the article are qualitatively identical using the inter-landmark distances d_i or the landmark Cartesian coordinates $\vec{\ell}_\alpha$ as facial vectors.

Separation of geometric and texture degrees of freedom. The face-space parametrisation is based, as previously mentioned, on the decoupling of *texture* (lightness, detailed, and skin textural) facial features, on the

one hand, and *geometric* (landmark coordinates), on the other hand. The separation of these two kinds of degrees of freedom is a standard paradigm of face representation (see, for example^{6,39,42}). It has been argued, in the light of the recently decoded neural coding for the facial identity in the primate brain, to be a naturally efficient parametrisation of the face⁴², outperforming other techniques in which texture and landmark-based are not separated, as the description in terms of *eigenfaces*.

Image deformation. Given a reference portrait (see Fig. 1B) and a vector of facial distances \mathbf{d}_1 , we create, by means of image (similarity transformation) deformation algorithms⁴⁷, a realistic facial image based on the reference portrait, deformed in such a way that the inter-landmark distances defined in Fig. 1A assume the desired values, $\mathbf{d} = \mathbf{d}_1$. Given the reference portrait image \mathcal{I}_0 , the position of its corresponding landmarks $\vec{c}_{0,\alpha}$, and the vector \mathbf{d} , we calculate the Cartesian coordinates $\vec{c}_{1,\alpha}$ of the new set of landmarks, completely defined by \mathbf{d} . The image deformation algorithm then generates a new facial image \mathcal{I}_1 with a point-dependent parameter linear transformation, such that the pixels occupying the landmark positions $\vec{c}_{0,\alpha}$ in the original image are mapped into the new positions $\vec{c}_{1,\alpha}$, and the rest of the pixels of the original image are mapped in order to produce a resulting image as realistic as possible. We have observed that, in order to produce realistic results, the linear transformation should be in the *similarity class*⁴⁷, beyond *affine* transformations. The deformed image is actually not created by mapping every pixel of the original image, but only the corners of a sub-grid; the sub-images inside each sub-grid are then warped to a polygon defined by the mapped corners of the grid, through affine transformations. The size of the sub-grid is taken to $\lesssim 15$ pixels. Both the reference portrait and the deformed images are roughly 300×400 pixels for RP1-2.

Genetic algorithm of face-space exploration. The genetic algorithm is based on a sequence of pairwise subject's choices among two facial images that are adaptively proposed to the subject, learned from his/her previous choices. An initial *population* of N vectors of randomised facial coordinates, $\mathbf{f}^{(s,m)}(0)$, evolve by means of genetic mutation and recombination, subject to the *selection* exerted by the experimental volunteer. At the t -th generation, the N vectors of the population generate an offspring of N individuals, by mutation and recombination according to the *differential evolution algorithm* (see Supplementary Sec. S3). The offspring is generated from the facial vectors only, independently of the reference portrait. The subject plays then the role of the evolutive pressure in the algorithm dynamics, selecting (N times) one among *two facial images*: one made from a vector of the population (and a reference portrait), and one made from its offspring. The $t + 1$ -th generation of vectors is then taken as the N vectors selected by the subject at the t -th generation. After a certain number, T , of generations, the population of facial vectors eventually reaches a regime in which the population of vectors do not change too much from one generation to the next. The T -th population of facial vectors is taken as the population of vectors sculpted by the subject, and constitutes the outcome of experiments E1-3.

This approach differs from previous approaches to facial attractiveness based on genetic algorithms^{48,49} in what: it allows a subject to select *in real time* a *realistic* facial image; in terms of geometric quantities only; with *fixed texture degrees of freedom*; finally, *avoiding the use of numerical ratings*, since the subject performs a sequence of left/right choices rather than assigning ratings to the images.

Populations of facial vectors sculpted by different subjects tend to be more far apart than populations sculpted by the same subject (see Sec. Results). Remarkably, the real difference between different subjects' attractors is even larger, since it is unavoidably underestimated in virtue of the finiteness of the experimental method. Indeed, two standard deviations with different origins contribute to the self-consistency distance μ_{sc} (see Fig. 3). One is the intrinsic, cognitive ambiguity of the subject's criterion; the other is the uncertainty brought by the genetic algorithm stochasticity (see Supplementary Sec. S3), whose origin is the discreteness of the proposed mutations and the consequent stochastic bias in the face space exploration. In genetic experiments with parameters in what we call in the *slow search regime* (mainly larger N and number of generations, T), the algorithmic uncertainty decreases, and μ_{sc} is expected to decrease consequently. This is the general expected behaviour of the differential evolution algorithm. We have also verified this fact experimentally: the distances among populations sculpted by a single subject significantly decrease for increasing values of $N = 10, 20, 28$. As a consequence, variants of the present experiment with *slower* genetic algorithm parameters would more finely *resolve* different subject's facial ideals, leading to a larger gap between inter-subject and self-consistency distances, at the cost of a larger number of subject's choices and experimental time.

Details of the experiments. Experiments E1, E2, E3 were performed by a pool of $S = 95$ volunteers (54 female, 39 male, of age average and standard deviation: 26(12)), mainly students, researchers and professors of the University "La Sapienza". Experiment E2 was performed under identical conditions of E1. A subset of $S_{sc} = 6$ participants to E1 (3 females, 3 males, of age average and standard deviation: 33(15)), were asked to perform 5 further instances of the experiment E1, in five different days, using, as in E1, the reference portrait RP1. The genetic algorithm parameters used are (see Supplementary Sec. S3): $N = 28$, $T = 10$, $\mu = 0.15$, $\rho = 1$. Each subject performed a number of $NT = 280$ choices among couples of facial images. These are 400×300 pixel, B/W images in an 1024×768 resolution monitor. The reference portraits RP1-2 have been taken from the *Chicago face database*⁵⁰. Each experiment lasted roughly 25 minutes on average (see the histogram of time intervals among successive left-right choices in figure Supplementary Sec. S7). The subjects were asked to look away and relax for some second each $N = 28$ choices. All methods in experiments E1-3 were carried out in accordance with relevant guidelines and regulations. The experimental protocols used have been approved by the General Data Protection Regulation (EU) 2016/679. Informed consent was obtained from all subjects. No subjects under 18 participated in the experiment.

Data Availability

The data and the codes devoted to the data analysis are available by request to the corresponding author.

References

- Bzdok, D. *et al.* ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Structure and Function* **215**, 209–223, <https://doi.org/10.1007/s00429-010-0287-4> (2011).
- Hahn, A. C. & Perrett, D. I. Neural and behavioral responses to attractiveness in adult and infant faces. *Neuroscience and Biobehavioral Reviews* **46**, 591–603, <https://doi.org/10.1016/j.neubiorev.2014.08.015>. Beyond Sexual Selection: the evolution of sex differences from brain to behavior (2014).
- Laurentini, A. & Bottino, A. Computer analysis of face beauty: A survey. *Computer Vision and Image Understanding* **125**, 184–199, <https://doi.org/10.1016/j.cviu.2014.04.006> (2014).
- Little, A. C. Facial attractiveness. *Wiley Interdisciplinary Reviews: Cognitive Science* **5**, 621–634, <https://doi.org/10.1002/wcs.1316> (2014).
- Thornhill, R. & Gangestad, S. W. Facial attractiveness. *Trends in Cognitive Sciences* **3**, 452–460, [https://doi.org/10.1016/S1364-6613\(99\)01403-5](https://doi.org/10.1016/S1364-6613(99)01403-5) (1999).
- Walker, M. & Vetter, T. Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of personality and social psychology* **110**, 609 (2016).
- Little, A., Jones, B. & DeBruine, L. Facial attractiveness: evolutionary based research. *Philosophical Transactions*, <https://doi.org/10.1098/rstb.2010.0404> (2011).
- Leopold, D. A. & Rhodes, G. A Comparative View of Face Perception. *Journal of comparative psychology (Washington, D.C.: 1983)* **124**, 233–251, <https://doi.org/10.1037/a0019460> (2010).
- Maret, S. M. & Harling, C. A. Cross-cultural perceptions of physical attractiveness: Ratings of photographs of whites by cruzans and americans. *Perceptual and Motor Skills* **60**, 163–166, <https://doi.org/10.2466/pms.1985.60.1.163> (1985).
- Langlois, J. H. *et al.* Maxims or myths of beauty? a meta-analytic and theoretical review. *Psychological bulletin* **126**, 390 (2000).
- Cunningham, M., Roberts, A., Barbee, A., Druen, P. & Wu, C.-H. Their ideas of beauty are, on the whole, the same as ours. *Journal of Personality and Social Psychology* **68**, <https://doi.org/10.1037/0022-3514.68.2.261> (1995).
- Bernstein, I. H., Lin, T.-D. & McClelland, P. Cross- vs. within-racial judgments of attractiveness. *Perception & Psychophysics* **32**, 495–503, <https://doi.org/10.3758/BF03204202> (1982).
- Thakerar, J. N. & Iwakaki, S. Cross-cultural comparisons in interpersonal attraction of females toward males. *The Journal of Social Psychology* **108**, 121–122, <https://doi.org/10.1080/00224545.1979.9711969> (1979).
- Perrett, D. I., May, K. A. & Yoshikawa, S. Facial shape and judgements of female attractiveness. *Nature* **368**, 239–242, <https://doi.org/10.1038/368239a0> (1994).
- Perret, D., Lee, K. & Penton-Voak, I. E. A. Effects of sexual dimorphism on facial attractiveness. *Nature*, <https://doi.org/10.1038/29772> (1998).
- Alley, T. R. & Cunningham, M. R. Averaged faces are attractive, but very attractive faces are not average. *Psychological Science*, <https://doi.org/10.1111/j.1467-9280.1991.tb00113.x> (1991).
- Valenzano, D. R., Mennucci, A., Tartarelli, G. & Cellerino, A. Shape analysis of female facial attractiveness. *Vision Research* **46**, 1282–1291, <https://doi.org/10.1016/j.visres.2005.10.024> (2006).
- Cunningham, M. R., Barbee, A. P. & Pike, C. L. What do women want? facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of personality and social psychology* **59**, 61 (1990).
- Edler, R. J. Background considerations to facial aesthetics. *Journal of Orthodontics* **28**, 159–168, <https://doi.org/10.1093/ortho/28.2.159> (2001).
- Valentine, T., Lewis, M. B. & Hills, P. J. Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology* **69**, 1996–2019, <https://doi.org/10.1080/17470218.2014.990392> (2016).
- Hill, H. *et al.* How different is different? criterion and sensitivity in face-space. *Frontiers in Psychology* **2**, 41, <https://doi.org/10.3389/fpsyg.2011.00041> (2011).
- Joy, K. L. & Primeaux, D. A Comparison of Two Contributive Analysis Methods Applied to an ANN Modeling Facial Attractiveness. In *Fourth International Conference on Software Engineering Research, Management and Applications (SERA'06)*, 82–86, <https://doi.org/10.1109/SERA.2006.2> (2006).
- Schmid, K., Marx, D. & Samal, A. Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios. *Pattern Recognition* **41**, 2710–2717, <https://doi.org/10.1016/j.patcog.2007.11.022> (2008).
- Pallett, P. M., Link, S. & Lee, K. New golden ratios for facial beauty. *Vision Research* **50**, 149–154, <https://doi.org/10.1016/j.visres.2009.11.003> (2010).
- Fan, J., Chau, K., Wan, X., Zhai, L. & Lau, E. Prediction of facial attractiveness from facial proportions. *Pattern Recognition* **45**, 2326–2334, <https://doi.org/10.1016/j.patcog.2011.11.024>, Brain Decoding (2012).
- Shen, H. *et al.* Brain responses to facial attractiveness induced by facial proportions: evidence from an fMRI study. *Scientific Reports* **6**, <https://doi.org/10.1038/srep35905> (2016).
- Gray, D., Yu, K., Xu, W. & Gong, Y. Predicting facial beauty without landmarks. *Computer Vision—ECCV* **2010**, 434–447 (2010).
- Bottino, A. & Laurentini, A. The Intrinsic Dimensionality of Attractiveness: A Study in Face Profiles. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, Lecture Notes in Computer Science, 59–66, https://doi.org/10.1007/978-3-642-33275-3_7 (Springer, Berlin, Heidelberg, 2012).
- Eisenhal, Y., Dror, G. & Ruppim, E. Facial attractiveness: Beauty and the machine. *Neural Computation* **18**, 119–142, <https://doi.org/10.1162/089976606774841602>, PMID: 16354383 (2006).
- Mu, Y. Computational facial attractiveness prediction by aesthetics-aware features. *Neurocomputing* **99**, 59–64, <https://doi.org/10.1016/j.neucom.2012.06.020> (2013).
- Gan, J., Li, L., Zhai, Y. & Liu, Y. Deep self-taught learning for facial beauty prediction. *Neurocomputing* **144**, 295–303, <https://doi.org/10.1016/j.neucom.2014.05.028> (2014).
- Bronstad, P. M., Langlois, J. H. & Russell, R. Computational Models of Facial Attractiveness Judgments. *Perception* **37**, 126–142, <https://doi.org/10.1068/p5805> (2008).
- Chen, Y., Mao, H. & Jin, L. A novel method for evaluating facial attractiveness. In *2010 International Conference on Audio, Language and Image Processing*, 1382–1386, <https://doi.org/10.1109/ICALIP.2010.5685007> (2010).
- Xu, J. *et al.* Facial attractiveness prediction using psychologically inspired convolutional neural network (PI-CNN). In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1657–1661, <https://doi.org/10.1109/ICASSP.2017.7952438> (2017).
- Galantucci, L. M., Gioia, E. D., Lavecchia, F. & Percoco, G. Is principal component analysis an effective tool to predict face attractiveness? a contribution based on real 3d faces of highly selected attractive women, scanned with stereophotogrammetry. *Medical & Biological Engineering & Computing* **52**, 475–489, <https://doi.org/10.1007/s11517-014-1148-8> (2014).
- Adolphs, R., Nummenmaa, L., Todorov, A. & Haxby, J. V. Data-driven approaches in the investigation of social perception. *Phil. Trans. R. Soc. B* **371**, 20150367, <https://doi.org/10.1098/rstb.2015.0367> (2016).

37. Oosterhof, N. N. & Todorov, A. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences* **105**, 11087–11092, <https://doi.org/10.1073/pnas.0805664105> (2008).
38. Todorov, A. & Oosterhof, N. N. Modeling Social Perception of Faces [Social Sciences]. *IEEE Signal Processing Magazine* **28**, 117–122, <https://doi.org/10.1109/MSP.2010.940006> (2011).
39. Abir, Y., Sklar, A. Y., Dotsch, R., Todorov, A. & Hassin, R. R. The determinants of consciousness of human faces. *Nature Human Behaviour* **1**, <https://doi.org/10.1038/s41562-017-0266-3> (2017).
40. Farkas, L. G. *Anthropometry of the Head and Face*. (Raven Pr, 1994).
41. Burton, A. M., Miller, P., Bruce, V., Hancock, P. J. B. & Henderson, Z. Human and automatic face recognition: a comparison across image formats. *Vision Research* **41**, 3185–3195, [https://doi.org/10.1016/S0042-6989\(01\)00186-9](https://doi.org/10.1016/S0042-6989(01)00186-9) (2001).
42. Chang, L. & Tsao, D. Y. The code for facial identity in the primate brain. *Cell* **169**, 1013–1028.e14, <https://doi.org/10.1016/j.cell.2017.05.011> (2017).
43. Ibáñez-Berganza, M., Lancia, G. L., Amico, A., Monechi, B. & Loreto, V. Unsupervised inference approach to human facial preference. (*to be submitted*) (2019).
44. Rhodes, G. The evolutionary psychology of facial beauty. *Annual Review of Psychology* **57**, 199–226, <https://doi.org/10.1146/annurev.psych.57.102904.190208>, PMID: 16097897 (2006).
45. Little, A. C., Burt, D. M. & Perrett, D. I. What is good is beautiful: Face preference reflects desired personality. *Personality and Individual Differences* **41**, 1107–1118, <https://doi.org/10.1016/j.paid.2006.04.015> (2006).
46. Wang, N., Gao, X., Tao, D., Yang, H. & Li, X. Facial feature point detection: A comprehensive survey. *Neurocomputing*, <https://doi.org/10.1016/j.neucom.2017.05.013> (2017).
47. Schaefer, S., McPhail, T. & Warren, J. Image deformation using moving least squares. *ACM Trans. Graph.* **25**, 533–540, <https://doi.org/10.1145/1141911.1141920> (2006).
48. Johnston, V. S. & Franklin, M. Is beauty in the eye of the beholder? *Ethology and Sociobiology* **14**, 183–199, [https://doi.org/10.1016/0162-3095\(93\)90005-3](https://doi.org/10.1016/0162-3095(93)90005-3) (1993).
49. Wong, B. J. F., Karmi, K., Devicic, Z., McLaren, C. E. & Chen, W.-P. Evolving attractive faces using morphing technology and a genetic algorithm: A new approach to determining ideal facial aesthetics. *The Laryngoscope*, <https://doi.org/10.1097/MLG.0b013e31816bf545> (2008).
50. Ma, D. S., Correll, J. & Wittenbrink, B. The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods* **47**, 1122–1135, <https://doi.org/10.3758/s13428-014-0532-5> (2015).

Acknowledgements

We acknowledge Andrea Giansanti, Davide Iannuzzi and Giovanni Pezzulo for suggestions and bibliographic highlights, Scott D. Shaeffer and William Schueller for technical and informatic advices, Andrea Gabrielli for mathematical suggestions, and Fernanda Pereira da Cruz Benetti for her help with the manuscript.

Author Contributions

M.I.-B. and V.L. devised the project. M.I.-B. devised the inference scheme and developed the FACEEXPLORE software of face exploration and the data analysis software. M.I.-B. and V.L. organised the experiments. A.A. and M.I.-B. performed the data analysis. M.I.-B. and V.L. wrote the article.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-44655-9>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019