Psychomorphospace—From Biology to Perception, and Back: Towards an Integrated Quantification of Facial Form Variation

Katrin Schaefer
Department of Anthropology
University of Vienna
Vienna, Austria
katrin.schaefer@univie.ac.at

Philipp Mitteroecker
Department of Theoretical Biology
University of Vienna
Vienna, Austria
philipp.mitteroecker@univie.ac.at

Bernhard Fink
Department of Sociobiology/Anthropology
University of Göttingen
Göttingen, Germany
bfink@gwdg.de

Fred L. Bookstein
Department of Anthropology
University of Vienna
Vienna, Austria
&
Department of Statistics
University of Washington
Seattle, WA, USA
flb@stat.washington.edu

Abstract

Several disciplines share an interest in the evolutionary selection pressures that shaped human physical functioning and appearance, psyche, and behavior. The methodologies invoked from the disciplines studying these domains are often based on different rhetorics, and hence may conflict. Progress in one field is thereby hampered from effective transfer to others. Topics at the intersection of anthropometry and psychometry, such as the impact of sexual selection on the hominin face, are a typical example. Since the underlying (evolutionary) theory explicitly places facial form in the middle of a causal chain as the mediating variable between biological causes and psychological effects, a particularly convenient conceptual and analytic scenario arises as follows. Modern morphometrics allows analysis of shape both “backwards” (by regressions on biology) and “forwards” (via predictions of psychology). The two computations are commensurate, hence the two kinds of effects can be compared and evaluated as directions in the same morphospace. We suggest translating the morphometric methodology of “Darwinian aesthetics” into this space, where psychological and other processes of interest can be coded commensurately. Such a translation permits researchers to relate the effects of biological processes on form to the perceptions of the same processes in one unified “psychomorphospace.”

Keywords

anthropometrics, attractiveness, Darwinian aesthetics, evolutionary psychology, facial shape, geometric morphometrics, masculinity, thin-plate spline, 2D:4D

August 5, 2009; accepted August 13, 2009

Biological Theory 4(1) 2009, 98–106.

Biological and evolutionary anthropologists, human ethologists, and evolutionary psychologists all share an interest in the evolutionary forces and constraints that shaped and continue to shape our minds, our behaviors, and our bodies inside and out. Traditionally, studies of these processes consider one domain at a time. The more the disciplines involved differ in their *explanatory rhetorics*, however, the narrower the methodology that results, preventing progress in one field from effective transfer to and adoption by others.

Among the processes sharing responsibility for observed variation in modern human facial form are sexual selection and social selection (e.g., Gangestad 2000; Grammer et al. 2003). On the assumption that mate choosers developed preferences for physical traits that honestly signal mate value, such studies typically reason by confirming associations of certain attributions with certain morphological features presumably related to developmental and hormonal status (Gangestad and Simpson 2000; Scheyd et al. 2008).

If biological explanations are going to be involved, any study of human physical appearance and its perception needs to accommodate descriptor vectors and prior predictors (instead of pure significance testing). One must represent the variation in people's evaluations in ways that suit biological reasoning about causes: what is needed is information on covariances of features, not just ratings of averages. We suggest a novel method, based on anthropological and biometric reasoning, which puts the biological cause of variation in human facial and bodily morphology back at the center of perception studies within the context of sexual selection. As the purpose of this article is not a review of the literature, we illustrate our points with a few examples drawn from the recent body of literature on perceived male attractiveness. Our aim is not to correct the intention of these studies but to improve their execution by taking advantage of the full power of recent developments in the mathematical understanding of morphology.

Sexual Dimorphism and Facial Prototyping

In their influential *Nature* paper of 1998, Perrett et al. reported a sexual dimorphism of facial attractiveness findings: both Caucasian and Japanese subjects preferred a feminine face shape to an average shape both within and between populations. As the details of their methods may be responsible for this pattern to some extent, we take a little space to set it out step by step. Specifically, the composite images of Perrett et al. were constructed by delineating and averaging a list of facial features of the faces sampled. These features were averaged into a "mean shape" that was then used to warp each component face; the resulting warps were averaged in turn to produce the prototype image (Tiddeman et al. 2001). In this protocol, randomly varying traits that do not differ across two groups will tend toward their averages, while traits that are significantly different between groups will be more clearly contrasting in composites (Perrett et al. 1994). Therefore, Perrett et al. reasoned, the method would be ideal for visualizing otherwise subtle structural differences between groups (e.g., a larger, more square-shaped jaw in the case of more masculine males) and for assessing whether groups of observers are sensitive to these differences (Boothroyd et al. 2008).

Since 1998 much has changed, especially the texture component in the construction of these shape and color prototypes. As Tiddeman et al. (2001) present the texture-enhanced transformation process, after the target prototype is warped into the new shape, one builds wavelet pyramids from these two images and calculates their magnitudes. After rescaling, the subject's pyramid is collapsed to give the new image, serving in turn as the basis for further manipulations.

The samples that give rise to these facial prototypes can vary greatly in their design. They might be "biological," such as groups by geographic origin, age, or sex, or they may arise as different fractions of a (self-)rating scale, or may be manufactured using the magnitude of some "vector differences" in positions of feature points between any two composites.

Definition and Perception of Male Facial Masculinity

Facial masculinity in particular has been assessed using three main methods of defining a masculinity–femininity dimension: sexual dimorphism, perceived masculinity, and pubertal development. Many researchers, including Perrett et al. (1998), Penton-Voak and Perrett (2000), and Cornwell et al. (2004), have used the differences in shape between prototype male and female faces to manufacture face stimuli varying the extent of sexually dimorphic traits. Adding a positive multiple of the female-to-male differences to an average should increase the masculinity of a subject's facial image, while adding a negative multiple should increase its femininity. This method has been criticized in that manipulating the appearance of face images using a continuum defined by male and female prototypes does not necessarily reflect changes to facial appearance that are due to the influence of testosterone (Meyer and Quong 1999; Swaddle and Reierson 2002).

The technique developed by Perrett et al. (1998) basically generated an "extreme masculine face shape" as a linear extrapolation of the differences between male and female average faces. Johnston et al. (2001) suggested that this approach might not be valid for the purpose of depicting a unitary biological process should the shapes assessed be the expression of bone growth, regulated as it is by complex interactions between growth hormone, androgens, and estrogen—the extrapolation has extropolated all three of these underlying components, not any one separately. Fine psychology is not necessarily commensurate with fine biology! Johnston et al. (2001) tried to remedy this by transforming shape and color along a three-part continuum defined by female and male student faces and male model faces perceived as particularly masculine. Contrary to Perrett et al. they found that women generally preferred the more masculine male faces, and perceived masculinity has been found to correlate positively with salivary testosterone levels (Penton-Voak and Chen 2004).

To address similar criticisms of Perrett et al.'s (1998) methods for manipulating masculinity in male faces (via linear extrapolation of the differences between male and female average faces), Swaddle and Reierson (2002) tried to mimic shape variation corresponding to varying levels of testosterone during puberty. They digitally manipulated male facial proportions according to shape changes during puberty (e.g., Enlow and Hans 1996) and had subjects rate attractiveness and dominance of the resulting continuum of images. These morphological operations influenced the perceived dominance of the face (the more testosterone, the more dominance) more than they influenced attractiveness.

Since then almost all significant studies aiming at understanding sex differences in facial shape and its perception used some variant of these methods (e.g., Feinberg et al. 2006; Potter and Corneille 2008). A recent paper by DeBruine et al. (2006) is a typical example of the use of transformation endpoints (for manipulating faces) that come from ratings, in this case from male faces rated particularly feminine or masculine. They tested their prediction by (linearly) correlating three different transformation endpoints (perceived masculinity, adult female and adult male composites, and boys and young men) with some additional ratings and items from self-report.

What can we infer from the results of these studies? These findings are intriguing but also equivocal. Moreover, the ambiguity in the reports on women's preferences for male faces cannot be explained through the rather subtle menstrual-cycle-dependent variation in female masculinity preference among male faces discovered later (Jones et al. 2008; Little et al. 2008; but see Peters et al. 2009). Although the studies reporting to find an effect of menstrual cycle alterations on women's mate preferences currently seem to outweigh the ones that do not, they do not provide a detailed description of the "masculinity" features that are associated with these preferences. Hence, the evidence on the causal relation between hormone levels and facial characteristics still remains speculative, as does the link between facial features and the persons' subjective response. From a biological perspective, two essential questions remain at this point: (1) which *biological* processes do the vectors used for capturing facial morphology actually represent, and (2) do these transformations correspond to any natural variation at all? It appears in retrospect that the methodology of all these studies was mainly psychometric and not biological. In order to do biometrics, the variability of the stimuli needs to be preserved, but average facial configurations do not do that. If we talk about causation, we need to address this challenge with a method more tightly tied to the ways that biologists reason and the ways they get information about data.

From the point of view of the authors of this essay, as long as the only representation of stimuli is via manipulation of averaged pictures, the association between perception and morphology remains too vague. In the best of these designs we can learn that transformed pictures correlate with some rating protocol, but we cannot discern which of the features in the faces triggered those responses. In a less effective setting, "known shape variation" is "directly tested" by digital manipulation of prototypes mimicking the shape effect of, for example, varying levels of testosterone through puberty, and raters select the images that appear most attractive or most physically dominant. Such "selection" designs have three major flaws: first, shapes are varied only on one geometric dimension, whereas in reality, they vary on indefinitely many; second, one still cannot tell which feature(s) are responsible for the rating; third, in spite of statistical significance, correlations are never high. We end up clueless about the actual biology of whatever pattern of cues is being exploited in rating the composite faces.

Tools and Applications From Geometric Morphometrics

We suggest a new methodological framework for the assessment of facial shape, its biology, and its perception. The literature we have reviewed focuses mainly on vague predictions of directions of preference or perceived gender; we need instead some quantitative method for assessing the cogency of the postulated biological process that causes perceptions. A quantitative state-of-the-art assessment of the morphology is missing.

Biological anthropology has always been concerned with the biological determinants of craniofacial and bodily shape. For about 10 years now, we have been applying geometric morphometric methods (GMM) to osteological and fossil material (e.g., Bookstein et al. 1999; O'Higgins and Collard 2002; Bookstein et al. 2003; Mitteroecker et al. 2004, 2005; Schaefer et al. 2004; O'Higgins et al. 2006; Schaefer et al. 2006b; Gunz et al. 2009) and also to facial soft tissue (Fink et al. 2005; Schaefer et al. 2005; Komori et al. 2009a,b). But the impact of facial shape on perception has only recently begun to be explored (Schaefer et al. 2006a; Windhager et al. 2008).

In all these applications, geometric morphometrics served for assessing and visualizing biological form and its covariates. The GMM toolkit is an adaptation of multivariate statistics and graphics to the study of phenotypic variation; it avoids some of the traditional algebraic pitfalls. In this method, the relative locations of a set of individually identified points—"landmarks" that are biologically homologous (preferably as points, and otherwise as bounding curves or surfaces)—are refashioned into a set of regular biometric variables, the shape coordinates, that can then be regressed one by one on the factors that cause them or the features of the systems they are presumed to affect. The methodological approaches of geometric morphometrics (Bookstein 1991; Rohlf and Marcus 1993; Marcus et al. 1996; Dryden and Mardia 1998; Slice 2005, 2007) make use of two- or three-dimensional coordinate data to describe size and shape at the same time, rather than using interlandmark distances or areas for size and, at the same time, angles and ratios for shape. The mathematical theorems and the biological axioms of geometric morphometrics are well understood (Bookstein 1991, 1996; Marcus et al. 1996; Dryden and Mardia 1998; Rohlf 1999; O'Higgins 2000; Slice 2005); its statistical properties have been proven superior to those of distance- or angle-based methods (Rohlf 2000, 2003) and supply graphics that are far more legible and interpretable by the applied biologist.

Procrustes Superposition, Procrustes Shape Coordinates, Procrustes Distance

Coordinate data vary among forms in terms of shape and scale but usually also differ by biologically uninformative features such as location and orientation of the forms at the time of digitizing. The first steps in comparing form between landmark configurations, therefore, require that these differences in the raw coordinates be minimized. The scale of the landmark configuration is represented by centroid size, defined as the square root of the sum of squared Euclidean distances from each landmark to the mean of the configuration of landmark coordinates (Bookstein 1991). Once the landmark coordinates have been scaled to unit centroid size, differences in the locations (translation) and orientations (rotation) of the scaled landmark configurations are then minimized using the generalized Procrustes analysis algorithm (GPA; Gower 1975; Rohlf and Slice 1990; Bookstein 1991; Goodall 1991; O'Higgins and Jones 1998). GPA centers and rotates all landmark coordinates relative to a tentative mean Procrustes shape so as to minimize the sum of squared distances between equivalent landmarks, and then recomputes that tentative mean shape so as to further reduce the same sum of squares. At the end of this process, the original coordinate data have been represented by "substitute Cartesian coordinates"—*shape coordinates*, as they vary around their own sample average shape. If shape variation in a sample is not too large, the scatter of points representing the shapes of the sample of landmark configurations can be projected into a linear Euclidean tangent space and then analyzed using standard multivariate methods (Dryden and Mardia 1993; Kent 1994; O'Higgins and Jones 1998; O'Higgins 2000).

Thin-Plate Spline and Visualizations

Unlike most biometric methods for which the passage from raw forms to decimal numbers is one-way, the GMM toolkit allows the biologist to return to the original space of laboratory bench or forest at the end of the analysis, so as to visualize statistical patterns as possible organismal shapes. The tool for this is the thin-plate spline (TPS; Bookstein 1991), which depicts any vector of the space of shape coordinates as a grid transformation. These TPS deformation grids illustrate changes of landmark configurations as deformations of the picture or the three-dimensional space in which they lie—as interpolation of the space between the landmarks.

Shape Regressions

As an example of the way conventional statistical ideas are modified to suit the GMM context, we explain the method of "shape regression" that generated the patterns shown in Figures 1 and 3. After the Procrustes steps just sketched, every landmark configuration has been assigned what is, in effect, a new set of Cartesian coordinates. If there were originally k landmarks, this is a total of $2k$ new coordinates: two for each of the points, just as at the time of original digitizing. The theorems of the Procrustes geometry allow the researcher to use these new coordinates just as if they were any other coherent set of variables—a list of scores on the items of some intelligence test, for instance, or the results of a standard medical blood test. If the task is to discover the shape pattern associated with a particular human response, such as a rating, we compute the ordinary regression of *each* of the shape coordinates on that rating, and then produce new shapes that correspond to the set of all the predicted coordinates (predicted value of the coordinate in each of the regressions in turn) for two extreme values of the rating along with the average rating. These predicted configurations are the configurations that are displayed in the figures, along with TPS deformation grids that help the comprehension of the pattern. If the task is to predict a shape from some measure of one of its biological causes, we carry out exactly the same regressions, one for each of the $2k$ shape coordinates on the claimed biological cause, and again we draw the predicted shapes for selected extreme values of that cause and also for the average value of the cause, where each predicted shape is accompanied by the corresponding TPS deformation grid from the average. For regressions of either type, statistical significance may be tested by a permutation test (Good 2000) of the total sum of explained variances over all of the shape coordinates: both x- and y-coordinates of all k landmark points.

How Does This Help Us With the Topic Under Interrogation?

In anthropology, these GMM methods have been primarily applied to questions concerning causes of human cranial form variation and the reconstruction of fossil crania (e.g., Bookstein et al. 1999; Strand Vidarsdottir et al. 2002; Gunz et al. 2009). They have also started to prove their usefulness in the study and explanation of living human form variation and of the perception of appearance and its relation to the forces that shaped it—the field bearing the charming name of "Darwinian aesthetics." It is now methodologically possible not just to isolate and plot the shape changes that are determined by proximate "biological causes" (Fink et al. 2005; Schaefer et al. 2006a,b; Schaefer and Bookstein in press) but also to work backward from perception to form—to map a rating (by a human observer of an image) back onto the image, thereby visualizing and then analyzing precisely those aspects of shape encoded in the rating (Schaefer et al. 2006a; Schaefer and Bookstein in press).

A particularly useful scenario arises when variables originating in one domain (for instance, auxology) are wielded to explain variation in another domain (e.g., behavior ratings) by a causal path running through actual organismal form. To summarize, this toolkit of GMM is a set of coherent statistical methods for the analysis of Cartesian coordinate data in causal chains like these. It is an adaptation of multivariate statistics and graphics to the study of phenotypic variation, and has already proved of high value for the detection of form changes that owe to such biological factors as growth, development, or hormones (e.g., O'Higgins 2000; Slice 2005). In this method, the relative locations of a set of individually identified points, or "landmarks," are encoded in a set of ordinary variables that can then be regressed one by one on the factors that cause them or features of the systems that they are presumed to affect. Corresponding to any such set of regression coefficients is that surprisingly clear diagram style, the "thin-plate spline" explained earlier, depicting their pattern in immediately interpretable terms as a grid transformation. When the underlying theory places organismal form in the middle of the causal chain, as the mediating variable between biological causes and psychological effects, the two sets of regression coefficients (one pointing "backwards" to biological causes, the other pointing "forwards" to psychological effects) can themselves be compared to see if they are aligned or not—to see if they convey one connected causal process. We herewith suggest exploiting this useful situation further by transferring the study of facial shape and its perception in toto to morphospace, where psychological and other processes of interest can be encoded via their correlations as displacement vectors.

Such a translation permits us to study and relate the effects of biological processes on form to the perceptions of the same processes in the same space, a fully function "psychomorphospace." Not only can we thereby compare the reality of the processes to the way they are perceived and the way they are expected to operate, but also we can simulate new forms and thereby test a wide range of current theories and speculations about the causes and consequences of variations of human form.

Biological Variables Affecting Facial Shape and Its Perception

An intriguing hypothesis is that intersexual selection has shaped the hominin face (Schaefer et al. 2004; Weston et al. 2004, 2007). As testosterone suppresses the immune system (Folstad and Karter 1992; Klein 2000), it has been suggested that "masculine" traits represent an honest signal of mate quality inasmuch as a male with high testosterone has successfully coped with its potential debilitating effects (Zahavi and Zahavi 1997; Kraemer 2000). On one theory, women should have evolved preferences for such masculine facial characteristics—should consider them attractive—because they signal immunocompetence and developmental or hormonal health (e.g., Symons 1995; Thornhill and Møller 1997; Thornhill and Grammer 1999). Several studies lend support to the view that women's preferences for facial "hormone markers" may reflect an index of mate quality (for a recent review see Grammer et al. 2003), but not all studies have this finding (e.g., Perrett et al. 1998).

To exemplify the methodological and conceptual framework that we outlined above, we cast the question of male facial variation and perception into a single Procrustes morphospace. We first try to clarify the morphological pattern corresponding to salivary testosterone, then prenatal testosterone, then the facial shape predicted by women's preference and masculinity assessment.

Regarding the biological causes under consideration, it is convenient to summarize our own earlier studies (e.g., Neave et al. 2003; Fink et al. 2005; Schaefer et al. 2005) on salivary testosterone and 2D:4D ratio versus male facial characteristics. The 2D:4D ratio (the relative length of the index [2D] to ring [4D] finger) is widely accepted as a proxy for early (fetal) testosterone/estrogen concentration (Manning et al. 1998; Lutchmaya et al. 2004). Facial photographs of 46 British adult men were characterized by 51 two-dimensional landmarks, and the resulting shape coordinates regressed upon the men's 2D:4D ratios and on their (circulating) salivary testosterone (Figures 1(a) and (b)). In neither shape regression does the variance explained exceed 5% (2.5% for 2D:4D, 2.1% for salivary testosterone). The two patterns that these independent variables predict turn out not to be particularly similar. The (general) expectation had been that increasing testosterone operates somewhat comparably on facial shape independent of the particular time of exposure onset. This does not seem to be the case. Some striking differences are at large scale (Figures 1(a) and (b)): low 2D:4D ratio results in a relatively robust and prominent lower face, whereas circulating testosterone seems to cause a rather uniform elongation of the face (Schaefer et al. 2005). If prenatal and pubertal testosterone has differential effects on male facial shape, this must be accommodated in future studies of women's preferences.

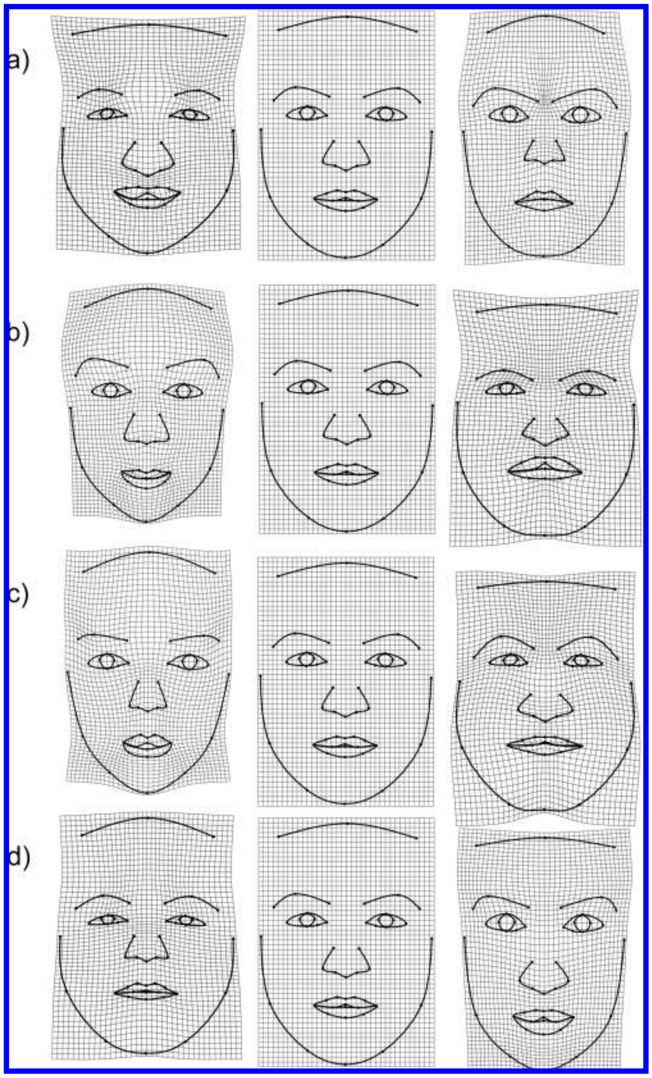

Figure 1.
Visualization of symmetrized shape regressions by TPS visualization grids. The middle panel, with the undeformed grid, is the average landmark configuration. Neighboring panels show the regression-based transformation in both directions from this consensus. The deformations correspond to a decrease (left) or an increase (right) on the independent variable: in (a) by 20 units on the salivary testosterone scale, (b) by 0.25 units on the prenatal testosterone level scale (2D:4D ratio), (c) by 7 units on the perceived masculinity scale (identical with the pattern for the perceived dominance scale), (d) by 7 units on the perceived attractiveness scale. The units are chosen for legibility of the warps as the actual sample variations are too small to depict the distortions effectively. Shape regressions did not reach statistical significance after 5000 permutations (Good 2000). See text for the argument that rows 2 and 3 are the most similar.

We further compare the shape pattern these biological factors determine with those used as a basis for masculinity (and dominance) and attractiveness perception in the very same individuals (original data published by Neave et al. 2003). For this, we visualize the shape regressions upon these two items comprising ratings by 36 women on one of two scales (Figures 1(c) and (d)). Comparing these two shape patterns associated with perception, it becomes clear that they are not the same but point at different facial features determining these attributions. When comparing them to the patterns determined by their putative biological causes, one finds some agreements but the regional similarities between the 2D:4D shape pattern and the one for perceived masculinity suggest a closer link between these two than among any other pair. This supports the claims of Neave et al. (2003) that some adult male facial characteristics perceived as dominant and masculine might be determined very early in ontogeny. It still remains to assess *how* similar or dissimilar these shape patterns are compared to each other, and the introduction of the combined morphospace allows the realization of this ambition.

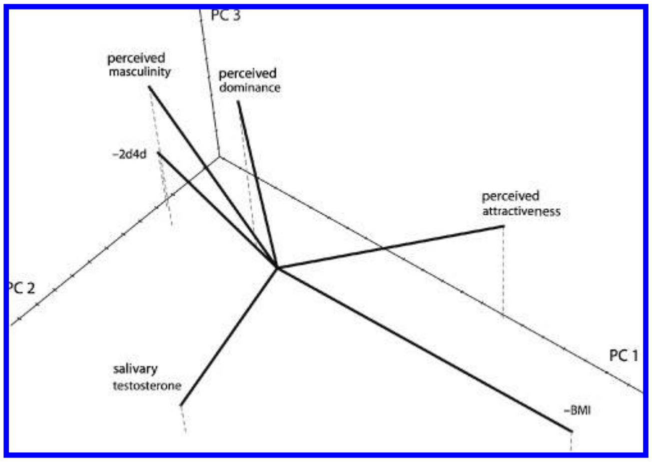

Figure 2.
Eigendecomposition of the shape vectors of the four regression coefficients visualized in Figure 1 (and of perceived dominance and BMI) in form space scaled by the shape effect of one standard deviation per predictor variable.

The Intersection of Anthropometry and Psychometry: Psychomorphospace for Shape Comparison

Evolutionary theory would explicitly place facial form in the middle of the causal chain, the mediating variable between biological causes and psychological effects. The different sets of shape regression coefficients (one set pointing "backwards" to a biological substrate, another "forwards" to perception) can themselves be compared and evaluated as to whether they are aligned. A singular value decomposition (SVD) ordinates the regression coefficients of interest (the one upon salivary testosterone, 2D:4D, perceived masculinity, dominance, and attractiveness, and also one upon body mass index (BMI), because generally body fat has turned out to considerably affect facial shape) as principal components of the corresponding shape vectors, each scaled with the shape effect of one standard deviation per predictor variable (raw data from Neave et al. 2003 and Schaefer et al. 2005). From these, we present the first three components (Figure 2). This visualization allows a novel possibility of assessing similarities in the pattern of the predictor variables originally extracted from biological

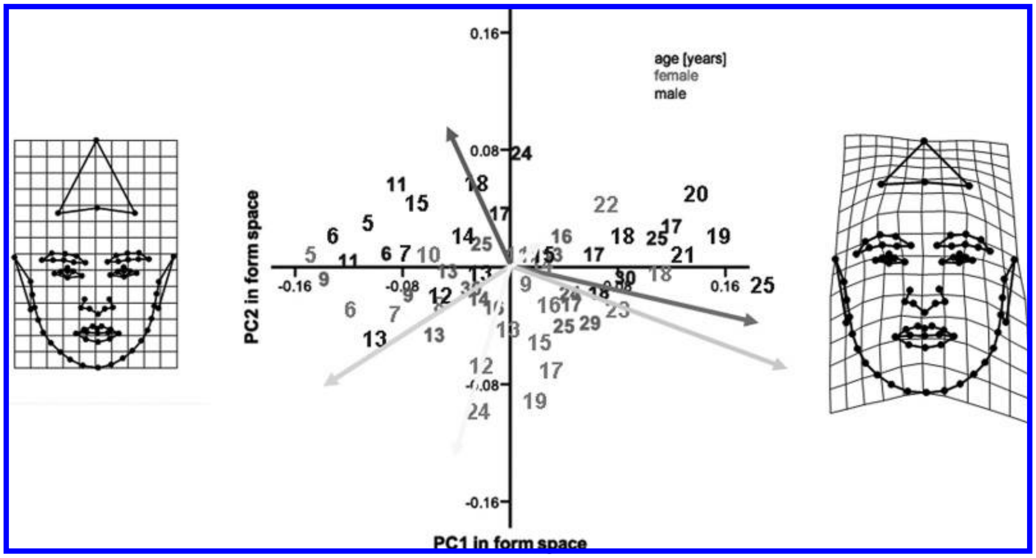

Figure 3.
Example for the general concept of the psychomorphospace: multiple causal analyses in one common vector space of shape features. Psychological as well as biological variables can be filled in as vectors (illustrated by the gray arrows in different shades) and shown together with actual shape variation (the numbers in the plot represent age in years corresponding to individual facial shapes).

measures and/or ratings (of the same body part)—whether they are aligned or not. The first dimension is clearly spanned by facial shapes corresponding to a low BMI on the one side, and a cluster of three other variables belonging to shapes that broadly relate to high testosterone, either measured or perceived. The second dimension is composed of the shape vector yielded by the attractiveness ratings and the one by salivary testosterone levels, pointing in opposite directions, and the third dimension might be interpreted as differentiating between biological and psychological variables obtained by ratings. This little graph with its six vectors is the first to allow for a direct comparison between shape properties determined by certain biological variables and their expected psychological effects.

A hypothesis that facial testosterone features are perceived as dominant and attractive can easily be checked here. In our small example, this presumption can be confronted much more directly than before. We find the two vectors relating to androgen influence on shape are not aligned but rather occupy different positions in the first as well as in the second component, indicating that 2D:4D and salivary testosterone operate differently on facial shape. The two vectors originating from the ratings upon perceived masculinity and attractiveness are roughly comparable in the first component but divergent in the second. TPS visualizations (Figure 1) of facial shape regressions upon perceived masculinity and dominance likewise do not resemble those upon salivary testosterone but rather those upon 2D:4D. Thus, aspects of male facial characteristics that convey information as relevant as dominance are determined as early as prenatally via a difference in androgen exposure.

When the methods are powerful enough, even domains as different as the biological causes and the psychological effects of different shape can be integrated for meaningful comparisons like these. Indeed, any methodology for evolutionary expectations about facial sexual dimorphism must rest on direct interplay between these two realms. For instance, someone may have predicted that the physique of men high in testosterone should signal immunocompetence and thus be perceived as attractive—we can test this by asking if the shape features pattern correlated with testosterone is the feature pattern that is perceived as attractive. A more subtle assumption about the hormonal state of the raters (Jones et al. 2008) drives a more detailed comparison in the form of an interaction term (ovulation by rating pattern), and so on. Many hypotheses about intersexual selection can be tested more powerfully by comparing the biologically predicted pattern of features to the pattern correlated best with the corresponding (biologically meaningful) rating.

This method suggests looking at a wide range of predictions pertinent to the field of Darwinian aesthetics. We can explore any shape pattern at the intersection between facial form and its perception whether or not the context is one of adaptive explanations. Figure 3 schematizes the general frame for this procedure. In psychomorphospace, the individual configurations are plotted in form space (Mitteroecker et al. 2004), the predictor variables of interest “filled in” as vectors, and their unique role realized as mediators between the two poles of an explicitly interdisciplinary theory. This way, it is possible not only to compare the shape pattern of a form change caused by any physical process and its effect on perception but also to systematically create virtual faces corresponding to any shape regression of interest and have them rated for the properties one might want to test.

In summary, the psychomorphospace concept introduced in this article places organismal form in the middle of the causal chain, as the mediating variable between biological causes and psychological effects. Its two vectors of regression coefficients (one pointing "backwards" to biology, the other pointing "forwards" to psychology) can themselves be compared to see if they are aligned or not—to see if they convey one connected causal process. Clearly, we favor transferring the study of Darwinian aesthetics in toto to this GMM morphospace, where psychological and other processes of interest can be encoded (via their cosines) as displacement vectors.

Acknowledgments

We are grateful to the Konrad Lorenz Institute for Evolution and Cognition Research for this unique workshop, to Werner Callebaut for all his editorial effort, and to Sonja Windhager for help with Figure 3. Likewise we are grateful to Nick Neave, whose data illustrated several of the main points of the article. We were supported partly by the DFG Emmy Noether Research Program "Evolutionary Psychology of Human Physical Appearance & Body Movement" (FI-1450/4–1) and by the EU FP6 Marie Curie Action Program EVAN MRTN-CT-2005–019564.

References

Bookstein FL (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. New York: Cambridge University Press.

Bookstein FL (1996) Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology* 58: 313–365.

Bookstein FL, Gunz P, Mitteroecker P, Prossinger H, Schaefer K, Seidler H (2003) Cranial integration in homo: Singular warps analysis of the midsagittal plane in ontogeny and evolution. *Journal of Human Evolution* 44: 167–187.

Bookstein FL, Schaefer K, Prossinger H, Seidler H, Fieder M, Stringer C, Weber GW, Arsuaga J-L, Slice DE, James Rohlf F, Recheis W, Mariam AJ, Marcus LF (1999) Comparing frontal cranial profiles in archaic and modern homo by morphometric analysis. *The Anatomical Record–The New Anatomist* 257: 217–224.

Boothroyd LG, Jones BC, Burt DM, DeBruine LM, Perrett DI (2008) Facial correlates of sociosexuality. *Evolution and Human Behavior* 29: 211–218.

Cornwell RE, Boothroyd L, Burt DM, Feinberg DR, Jones BC, Little AC, Pitman R, Whiten S, Perrett DI (2004) Concordant preferences for opposite-sex signals? Human pheromones and facial characteristics. *Proceedings of the Royal Society B: Biological Sciences* 271: 635–640.

DeBruine LM, Jones BC, Little AC, Boothroyd LG, Perrett DI, Penton-Voak IS, Cooper PA, Penke L, Feinberg DR, Tiddeman BP (2006) Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences* 273: 1355–1360.

Dryden IL, Mardia KV (1993) Multivariate shape analysis. *Sankya* 55: 460–480.

Dryden IL, Mardia KV (1998) *Statistical Shape Analysis*. Chichester, UK: Wiley.

Enlow DH, Hans MG (1996) *Essentials of Facial Growth*. Philadelphia: Saunders.

Feinberg DR, Jones BC, Law Smith MJ, Moore FR, DeBruine LM, Cornwell RE, Hillier SG, Perrett DI (2006) Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior* 49: 215–222.

Fink B, Grammer K, Mitteroecker P, Gunz P, Schaefer K, Bookstein FL, Manning JT (2005) Second to fourth digit ratio and face shape. *Proceedings of the Royal Society B* 272: 1995–2001.

Folstad I, Karter AJ (1992) Parasites, bright males, and the immunocompetence handicap. *American Naturalist* 139: 603–622.

Gangestad SW (2000) Human sexual selection, good genes, and special design. *Annals of the New York Academy of the Sciences* 907: 50–61.

Gangestad SW, Simpson JA (2000) On the evolutionary psychology of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences* 23: 573–587.

Good P (2000) *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. New York: Springer.

Goodall CR (1991) Procrustes methods and the statistical analysis of shape. *Journal of the Royal Statistical Society B* 53: 285–340.

Gower JC (1975) Generalized Procrustes analysis. *Psychometrika* 40: 33–51.

Grammer K, Fink B, Møller AP, Thornhill R (2003) Darwinian aesthetics: Sexual selection and the biology of beauty. *Biological Reviews of the Cambridge Philosophical Society* 78: 385–407.

Gunz P, Bookstein FL, Mitteroecker P, Stadlmayr A, Seidler H, Weber GW (2009) Early modern human diversity suggests subdivided population structure and a complex out-of-Africa scenario. *Proceedings of the National Academy of Sciences USA* 106: 6094–6098.

Johnston VS, Hagel R, Franklin M, Fink B, Grammer K (2001) Male facial attractiveness: Evidence for hormone-mediated adaptive design. *Evolution and Human Behavior* 22: 251–267.

Jones BC, DeBruine LM, Perrett DI, Little AC, Feinberg DR, Law Smith MJ (2008) Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior* 37: 78–84.

Kent JT (1994) The complex Bingham distribution and shape analysis. *Journal of the Royal Statistical Society B* 56: 285–299.

Klein SL (2000) The effects of hormones on sex differences in infection: From genes to behavior. *Neuroscience and Biobehavioral Reviews* 4: 627–638.

Komori M, Kawamura S, Ishihara S (2009a) Averageness or symmetry: Which is more important for facial attractiveness? *Acta Psychologica* 131: 136–140.

Komori M, Kawamura S, Ishihara S (2009b) Effect of averageness and sexual dimorphism on the judgment of facial attractiveness. *Vision Research* 49: 862–869.

Kraemer S (2000) The fragile male. *British Medical Journal* 321: 1609–1612.

Little AC, Jones BC, DeBruine LM (2008) Preferences for variation in masculinity in real male faces change across the menstrual cycle: Women prefer more masculine faces when they are more fertile. *Personality and Individual Differences* 45: 478–482.

Lutchmaya S, Baron-Cohen S, Raggatt P, Knickmeyer R, Manning JT (2004) 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development* 77: 23–28.

Manning JT, Scutt D, Wilson J, Lewis-Jones DI (1998) The ratio of 2nd to 4th digit length: A predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction* 13: 3000–3004.

Marcus LF, Corti M, Loy A, Naylor GJP, Slice DE (1996) *Advances in Morphometrics*. New York: Plenum Press.

Meyer DA, Quong MW (1999) The bio-logic of facial geometry. *Nature* 397: 661–662.

Mitteroecker P, Gunz P, Bernhard M, Schaefer K, Bookstein FL (2004) Comparison of cranial ontogenetic trajectories among great apes and humans. *Journal of Human Evolution* 46: 679–698.

Mitteroecker P, Gunz P, Bookstein FL (2005) Heterochrony and geometric morphometrics: A comparison of cranial growth in *Pan paniscus* versus *Pan troglodytes*. *Evolution and Development* 7: 244–258.

Neave N, Laing S, Fink B, Manning JT (2003) Second to fourth digit ratio, testosterone and perceived male dominance. *Proceedings of the Royal Society B* 270: 2167–2172.

O'Higgins P (2000) The study of morphological variation in the hominid fossil record: Biology, landmarks and geometry. *Journal of Anatomy* 197: 103–120.

O'Higgins P, Bastir M, Kupczik K (2006) Shaping the human face. *International Congress Series* 1296: 55–73.

O'Higgins P, Collard M (2002) Sexual dimorphism and facial growth in papionin monkeys. *Journal of Zoology* 257: 255–272.

O'Higgins P, Jones N (1998) Facial growth in *Cercocebus torquatus*: An application of three-dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy* 193: 251–272.

Penton-Voak IS, Chen JY (2004) High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior* 25: 229–241.

Penton-Voak IS, Perrett DI (2000) Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behavior* 21: 39–48.

Perrett DI, Lee KJ, Penton-Voak I, Rowland D, Yoshikawa S, Burt DM, Henzi SP, Castles DL, Akamatsu S (1998) Effects of sexual dimorphism on facial attractiveness. *Nature* 394: 884–887.

Perrett DI, May KA, Yoshikowa S (1994) Facial shape and judgements of female attractiveness. *Nature* 368: 239–242.

Peters M, Simmons LW, Rhodes G (2009) Preferences across the menstrual cycle for masculinity and symmetry in photographs of male faces and bodies. *PLoS ONE* 4(1) e4138.

Potter T, Corneille O (2008) Locating attractiveness in the face space: Faces are more attractive when closer to their group prototype. *Psychonomic Bulletin and Review* 15: 615–622.

Rohlf FJ (1999) Shape statistics: Procrustes superimpositions and tangent spaces. *Journal of Classification* 16: 197–223.

Rohlf FJ (2000) Statistical power comparisons among alternative morphometric methods. *American Journal of Physical Anthropology* 111: 463–478.

Rohlf FJ (2003) Bias and error in estimates of mean shape in geometric morphometrics. *Journal of Human Evolution* 44: 665–683.

Rohlf FJ, Marcus LF (1993) A revolution in morphometrics. *Trends in Ecology and Evolution* 8: 129–132.

Rohlf FJ, Slice DE (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* 39: 40–59.

Schaefer K, Bookstein FL (in press) Does geometric morphometrics serve the needs of plasticity research? *Journal of Biosciences* 34(1).

Schaefer K, Fink B, Grammer K, Mitteroecker P, Gunz P, Bookstein FL (2006a) Female appearance: Facial and bodily attractiveness as shape. *Psychological Science* 48: 187–204.

Schaefer K, Fink B, Mitteroecker P, Neave N, Bookstein FL (2005) Visualizing facial shape regression upon 2nd to 4th digit ratio and testosterone. *Collegium Antropologicum* 29: 415–419.

Schaefer K, Lauc T, Mitteroecker P, Gunz P, Bookstein FL (2006b) Dental arch asymmetry in an isolated Adriatic community. *American Journal of Physical Anthropology* 129: 132–142.

Schaefer K, Mitteroecker P, Gunz P, Bernhard M, Bookstein FL (2004) Craniofacial sexual dimorphism patterns and allometry among extant hominids. *Annals of Anatomy* 186: 471–478.

Scheyd GJ, Garver-Apgar CE, Gangestad SW (2008) Physical attractiveness: Signals of phenotypic quality and beyond. In: *Foundations of Evolutionary Psychology* (Crawford C, Krebs D, eds), 239–259. Hillsdale, NJ: Erlbaum.

Slice DE (2005) Modern morphometrics. In: *Modern Morphometrics in Physical Anthropology* (Slice DE, ed), 1–45. New York: Kluwer/Plenum.

Slice DE (2007) Geometric morphometrics. *Annual Review of Anthropology* 36: 261–281.

Strand Vidarsdottir U, O'Higgins P, Stringer C (2002) A geometric morphometric study of regional differences in the ontogeny of the modern human facial skeleton. *Journal of Anatomy* 201: 211–229.

Swaddle JP, Reierson GW (2002) Testosterone increases perceived dominance but not attractiveness in human males. *Education Economics* 269: 2285–2289.

Symons D (1995) Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness. In: *Sexual Nature, Sexual Culture* (Abramson PR, Pinkerton SD, eds), 80–120. Chicago: University of Chicago Press.

Thornhill R, Grammer K (1999) The body and face of woman: One ornament that signals quality? *Evolution and Human Behavior* 20: 105–120.

Thornhill R, Møller AP (1997) Developmental stability, disease and medicine. *Biological Reviews of the Cambridge Philosophical Society* 72: 497–548.

Tiddeman B, Burt M, Perrett D (2001) Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications* 21(5): 42–50.

Weston EM, Friday AE, Johnstone RA, Schrenk F (2004) Wide faces or large canines? The attractive versus the aggressive primate. *Proceedings of the Royal Society B* 271(S6): 416–419.

Weston EM, Friday AA, Liò P (2007) Biometric evidence that sexual selection has shaped the hominin face. *PloS ONE* 2(6): e710.

Windhager S, Slice DE, Schaefer K, Oberzaucher E, Thorstensen T, Grammer K (2008) Face to face: The perception of automotive designs. *Human Nature* 19: 331–346.

Zahavi A, Zahavi A (1997) *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford: Oxford University Press.