
Does geometric morphometrics serve the needs of plasticity research?

KATRIN SCHAEFER^{1,*} and FRED L BOOKSTEIN^{1,2}

¹*Department of Anthropology, University of Vienna, Vienna, Austria*

²*Department of Statistics, University of Washington, Seattle, USA*

*Corresponding author (Fax, +43-1-4277 9547; Email, katrin.schaefer@univie.ac.at)

The study of human craniofacial variation exemplifies general problems associated with the analysis of morphological plasticity that owe to the dependence of results on the methods by which phenotypic variation is quantified. We suggest a definition of plasticity that does not subordinate the developmental to the evolutionary: A process model in which changes are not a function of any mean or average, but only of the current state. Geometric morphometrics, a toolkit for assessing and visualizing biological form and its covariates, avoids some of the traditional pitfalls by focusing directly on the analysis of the two- and three-dimensional coordinates of anatomical landmarks. We discuss its potential relevance to phenotypic and developmental plasticity research, as well as some of its limitations, and demonstrate two useful analyses: assessment of asymmetry, and appraisal of integration. We itemize some of our previous studies on causes (inbreeding, environmental circumstances, etc.) and consequences (attractiveness perception) of asymmetry in humans, present some findings relating to the impact of sex on shape, and speculate about the adaptive relevance of one of these processes in particular. A closing argument points out that such considerations are possible only because of the careful separation of assumptions from empirical evidence entailed in the course of this type of data analysis.

[Schaefer K and Bookstein F L 2009 Does geometric morphometrics serve the needs of plasticity research?; *J. Biosci.* **34** 00–00]

1. Introduction: What are we doing in this Special Issue on *Phenotypic and Developmental Plasticity*?

We are interested in the organismal phenotype and the quantitative assessment of its variations of form. Even though the anthropologist who talks about differences in form refers mainly to human or primate shape variation, the methodological concerns are worth sharing more widely with the general readership of this *Journal of Biosciences*, and the results themselves may likewise be of broader interest. Thus our contribution combines some methodological tools with some applications from anthropology concerning the biology and perception of human craniofacial form and function, particularly the effects of stress and sex on shape.

Anthropologists cannot study phenotypic plasticity directly. On the other hand, the principal properties studied are not specific to primates—we deal with the responses of

a range of organisms to factors, responses often expressed as incremental (gradual) changes of size or shape. Well-controlled developmental and environmental plasticity studies can focus directly on the developmental reaction norm, but we must instead attend to the broader patterns that arise across samples as they vary by taxon, evolutionary time, individual age, sex, hormonal level, environmental circumstances, or the like. We are also interested in 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 these “biological causes” 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 analysing precisely those aspects of shape encoded in the rating.

Keywords. Asymmetry; craniofacial variation; geometric morphometrics; integration; plasticity; physical attractiveness; sexual dimorphism

Abbreviations used: DA, directional asymmetry; FA, fluctuating asymmetry; GMM, geometric morphometrics; GPA, Generalized Procrustes Analysis algorithm; PCA, principal components analysis; PCs, principal components; TPS, thin-plate spline

Despite the differences of purpose and style between anthropological studies and quantitative evo-devo studies, they share a preference for observing and modelling gradual changes in morphological characters, the observations typically associated with the statistical methods we call “biometrical”. Both kinds of study consequently rest on methodological decisions about how such gradual changes ought to be quantified. As scientists we all begin with a data set of phenotypes; whether we map them onto genotype and environment or onto other variables, the methods required are generally the same.

With these considerations in mind, our aim for this paper is threefold. First, we will briefly describe the geometric morphometrics (GMM) toolkit with an eye toward its potential utility in phenotypic plasticity research. Second, we will discuss the potential strengths and the very real limitations of these methods. Finally, we sketch results and implications of some of our own studies on the morphology of the human face that we think will be relevant to audiences interested in the topic of plasticity more broadly.

2. Presentation of some GMM tools, and explanation of basics

This very brief summary is elaborated, of course, in technical papers and textbooks to which we can only point in citations. The GMM toolkit, an adaptation of multivariate statistics and graphics to the study of phenotypic variation, has already proved to be quite useful for the detection of form changes that owe to such biological factors as growth, development, or hormones. 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 ordinary 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 (Rohlf and Bookstein 1990; Marcus *et al* 1996; Dryden and Mardia 1998; Slice 2005, 2007) make use of two-dimensional 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; O’Higgins 2000b; Slice 2005), and its statistical properties have been proven superior to those of distance-based or angle-based methods (Rohlf 2000a,b, 2003) and supply graphics that are far more legible and interpretable to the applied biologist.

2.1 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, scales, 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, which is mathematically unique and elegant, the original coordinate data have been represented by “substitute Cartesian coordinates”, *shape coordinates*, as they vary around their own sample average shape. (“Substitute Cartesian coordinates,” Bookstein 2009, is an interpretation of the Procrustes superposition as if each specimen had been digitized on a slightly different digitizer.) If shape variation in a sample is not too large, the scatter of points in shape space representing the shapes of the sample of landmark configurations can be projected into a linear, Euclidean tangent space and then analysed using standard multivariate methods (Dryden and Mardia 1993; Kent 1994; O’Higgins and Jones 1998; O’Higgins 2000a,b).

2.1.1 *Principal components*: The most commonly used multivariate method for exploring the shape variability of landmark configurations after GPA and tangent projection is principal components analysis (PCA; Smith 2002; Wikipedia 2008). A PCA carried out on the covariance matrix of the Procrustes shape coordinates results in a series of orthogonal linear combinations of the original variables (principal components: PCs), such that the first principal component (PC1) is a linear combination of the original data that accounts for the maximum possible variance. Subsequent PCs are constructed to account for the maximal remaining variance while remaining orthogonal to preceding PCs. The scores of shapes on PCs can be plotted to examine groupings by shape similarity, and the formulas for those linear combinations can be interpreted as deformation diagrams with the aid of the thin-plate spline

to be introduced presently. There is a version of multivariate statistics, *principal coordinates analysis*, in which these PC scores arise directly from the sums of squares of those forms at the time of computing the GPA (Gower 1966). (Thus, Principal components of Procrustes shape coordinates can be computed directly from Procrustes distances without actually producing the coordinates themselves.) At root, then, they reflect the scientist's choice of that minimizing sum of squares as a measure of shape-to-shape dissimilarity, the celebrated *Procrustes shape distance*. The principal component scores supply the graphic display within which distances on the page most closely match the actual Procrustes shape distances among the original digitized forms. Other choices of distance are certainly possible, but do not have the elegant mathematics of the Procrustes approach.

2.1.2 TPS 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 that does this, the thin-plate spline (TPS, Bookstein 1991), depicts any vector of the space of shape coordinates as a grid transformation of the sort invented by Albrecht Dürer (the German Renaissance painter, printmaker and theorist) and made familiar by the British natural philosopher D'Arcy Thompson (1917). TPS deformation grids depict changes of landmark configurations as deformations of the picture or the 3D space in which they lie—as interpolations of the space between the landmarks.

The assumptions underlying the TPS align well with the assumptions of biometrics as well as of the Procrustes analysis itself (Bookstein 1991). But please note that we are not taking a position regarding the actual core of current biometry: we are not assuming anything about the validity of these seemingly routine applications of ordinary statistical methods to the full range of biological questions. It is well-known, for instance, that the simplest statistics—averages and their comparisons—require a wide range of assumptions, of which several (such as the internal homogeneity or “exchangeability” of sampled entities) are rarely capable of empirical confirmation. The dilemmas of averages extend to the dilemmas of morphometrics because the fundamental underlying statistical construction of morphometrics, the regression coefficient, is itself an average (rather than any sort of scientific model in its own right). A *simple* regression coefficient does not necessarily stand in for a model; it has a much simpler interpretation, the weighted average of a ratio. (The usual regression coefficient for y on x is

$$\frac{\sum_1^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^n (x_i - \bar{x})^2} = \sum_1^n \left[\frac{y_i - \bar{y}}{x_i - \bar{x}} \frac{(x_i - \bar{x})^2}{\sum_1^n (x_i - \bar{x})^2} \right]$$

where the ratio being averaged is

$$\left(\frac{y_i - \bar{y}}{x_i - \bar{x}} \right)$$

and the weights are

$$\frac{(x_i - \bar{x})^2}{\sum_1^n (x_i - \bar{x})^2}$$

which sum to 1. This observation is originally due to Edgeworth in 1892.)

Morphometrics imports from multivariate biometrics some additional tools dealing with the handling of quantities like these in large groups—matrix inversion and diagonalization, singular-value decompositions—that make no reference in themselves to concepts of theoretical or mathematical biology and so will not be critiqued here. The point, rather, is that any invocation of multivariate statistics is in a context of samples from an unchanging distribution having means and covariances, i.e. averages of values and averages of ratios, and in a world of plastic phenomena, the underlying averages are not guaranteed to exist. We will see in section 5 some examples of studies in which plastic aspects of the history have been brought under strong numerical control, but first, in section 3, we need to sketch the troubled history of interactions between biometric and developmental models that accounts for the relative novelty of our approach within the world of evo-devo studies.

3. Limitations of geometric morphometrics as a methodology for studies of plasticity

Walter Elsasser, in his magnificent book-length essay *Reflections on a Theory of Organisms* (1998), draws a central contrast between the three core notions of the physical sciences—space, time, causality—and the four that underlie the organismal sciences: not only the naturalistic three, but also “memory”. Elsasser here explicitly meant non-material memory, coding memory—the way a fertilized egg “remembers” the pattern of the zygote that was its mother—but it is enough for our discussion that there is some memory of an actual developmental trajectory, an actual contingent history coded in an organism's physiology and its responses to novel challenges. Morphometrics has no formalisms for that memory, and borrowing from longitudinal data analysis will not resolve the problem, either, for the permanent shifts that result from a continued series of shocks such as characterize either evolution or development.

A dictionary informs us that the word *plasticity* means the property of responding to load with a permanent change of form. In this it differs from *elasticity*, which is a temporary response. These constructs differ radically in the way they can be articulated to morphometrics. The equations of

elasticity have to do with displacements from an equilibrium form, displacements that are likely to be linear in various matrices expressing the explicit forces of an experimental situation. The return to equilibrium, in other words, is in analogy to the role of the grand mean in the multivariate statistical formalisms that underlie GMM. In contrast, a plastic system, permanently deforming, loses the memory of its earlier state or its earlier developmental targets in the sheer factuality of its present condition. Neither GMM in particular nor biometrics in general has any way of encoding this path-dependence. Plastic development, properly construed, is not a suitable class of investigations for statistical study. Morphometrics embraces not the plastic but the elastic, in which the memory of the ideal state is permanent.

This sense of the word “plasticity” diverges from the sense that was more commonly relied on at the meeting from which this paper derived. Ordinarily, phenotypic plasticity is defined as the way in which the morphological expression of a given genotype varies under different environmental conditions (Encyclopædia Britannica 2008). It thus refers to “non-genetic variation”, that is to say, the effect of environment holding genotype constant. But genotype is no more “constant” than development—on time scales larger than that of the individual life span both are always changing, and we are interested in a method for studying evo-devo that treats them symmetrically rather than subordinating development to evolution the way the standard definition does. The definition of plasticity explored in this contribution achieves that symmetry by erasing any claims about a mean or average form from both types of explanation; both evolution and development are being modelled as processes without a memory any further back than the immediately previous state. For instance, this construal of plasticity, along with the usefulness of morphometrics in studies of plastic phenomena, is consistent with Benedikt Hallgrímsson’s applications to experiments with knock-out mice; such mice obviously develop but have no genetic history. Likewise, the terms “history” and “memory” are used here as Elsasser and Gaddis were already using them, in the sense of the physical natural sciences. Memory is the encoding of a system state no longer structurally present, such as a population mean, and history is a structural record of previous physical states.

The American historian John Lewis Gaddis took time off from his lifetime of studies of the Cold War between the United States and the Soviet Union to write a fine short book, *The Landscape of History* (2004), on this relation between science and narrative as seen from the other end. He noted that historical sciences, should they indulge in quantification at all, gain hardly anything from the quantitative style that imitates the social sciences. There the word “statistics” is grammatically plural. Statistics are applied in the original sense of the word, numbers relevant to the activities of a

nation-state along with the corresponding econometric, demographic, and “cliometric” formulas. No, Gaddis insisted, quantitative historical methodology is of value only when it imitates the methods of the *natural sciences* and out of those mainly the methodologies that are path-dependent: chaos theory, cosmology, plate tectonics, evolutionary hazard. There is a deep and possibly unintended irony here, as the statistics of the natural sciences actually originated in the social sciences, but that is not our point: rather, that the historical sciences align with the parts of the natural sciences *that are not statistical*.

In plasticity phenomena, a form forgets its origins but remembers its immediate past. The statistical methods suited to such studies thus are unlikely to be the same as the statistical methods arising in fields for entities that have no history, entities like random samples from an unchanging multivariate Gaussian distribution. But these Gaussians are the basic constructs of GMM models, as they have been the basic formalism of biometrical models ever since Francis Galton intuited the idea of regression in 1877.

Many infelicities follow from this fundamental antinomy. Geometric morphometrics presumes a *pointwise homology function*, a diffeomorphism (differentiable map with differentiable inverse) linking every pair of forms of a data set. But the essence of development is the innovation of form, or else its rearrangement in the course of metamorphosis; and the essence of evolutionary innovation is the reassignment of parts to other parts, or other functions, under the control of processes that are exquisitely specific not only in space (the Hox genes) but also in time (a channel of quantification that goes unformalized in GMM). These diffeomorphisms, then, are unlikely to correspond at all closely to biomathematical forms of explanation, whether of the biomechanical or the evo-devo framework.

Likewise, geometric morphometrics cannot truthfully model *influences on the organism*, both because it has no way of annotating the timing of those influences (onset, offset, time-course of amplitude) and because the principal such “influences” are wholly invisible to us as geometric statisticians—they arise instead in the genome, proteome, and metabolome that produce our observed organic forms as artifacts, afterthoughts. Morphometrics likewise has no room for the actual equations of biomechanics, physiology, and the other extant fully quantified biomedical sciences. The integrals that make up bioenergetics are not derived from the form of organs, but from their extent, which morphometrics ordinarily treats as epiphenomena if it does not ignore them entirely. Morphometrics has no comfortable role for cycles (gait, the cardiac cycle, ...), as distinct from studies of things holding still (e.g. the skulls of dead animals), nor any language, really, for the bifurcations at which organisms choose one path or another (male, female) over the course of development. An exception relevant to the

readers of this special issue is that environmental influences, when studied in the short term, might have a mean value; but in the long term, the environment is as plastic as anything else. In short, the language of morphometrics is no match for the language of plasticity in developmental theory. Its formalisms of homology, causation, and memory are all hopelessly inadequate.

And yet geometric morphometrics offers the best available tools for the numerical aggregation of experimental evidence about plasticity in its “frozen” form, images of the form achieved under controlled circumstances.

4. Morphometrics for plasticity research

There is no contradiction between this proposition and the critique of the preceding section. While morphometrics offers no language to match the theorist’s *explanations* of plastic phenomena, it has arrived at tools of immense elegance and power for the pattern analysis of subtle distinctions among the *outcomes* of these processes in settings characterized by controlled measurement. Morphometrics is a tool for studies of *small changes in outcome*, in systems for which the “plastic” part of the topic of plasticity is mainly, or wholly, metaphor. It is a method particularly suited to the theory-free parts of historical organismal biology: the parts *not* governed by equations dealing with energy, or force, nor the real random processes of thermodynamics or genetics. As a truly *statistical* method, GMM benefits from the machinery that underlies this entire branch of applied mathematics: clarity about empirical inference, about probabilities of hypotheses given the data, and about the actual limits of uncertainty (of value, of causation, or epistemology) in the particular situation under study. Morphometrics’ articulation with the actual theoretical concerns of the biological sciences depends on the skill of the investigator in framing the study sample, the interactions with it and the measurements that archive it: but that is all that one can reasonably expect from *any* quantitative method, statistical-mathematical or otherwise.

Benedikt Hallgrímsson and colleagues (2002) portray the study of phenotypic variability as directed toward explaining three related patterns of phenotypic variation: (i) variation in the sensitivity of developmental systems to genetic or environmental perturbation, or “canalization”; (ii) variation in how predictable development is under the same environmental and genetic conditions, or “developmental stability”; and (iii) the tendency for multivariate variation to be structured by developmental or functional relationships, “morphological or developmental integration”. There is room only to sketch briefly two of the core GMM methods which we think are applicable here. One of the core topics, the study of *symmetry*, has a history in art and in technology of more than two thousand years (Weyl 1952) but arises

here by virtue of the usefulness of measures of asymmetry in studies of developmental instability (Ludwig 1932; Polak 2003). Morphometrics instantiates this ancient praxis in the guise of biometrics (vectors, sums of squares) and in so doing greatly extends the empirical power of the old language. In contrast, the other topic, integration, appears in today’s biological theory principally as a seductive metaphor (Pigliucci and Preston 2004). Although its instantiations in earlier cycles of biometric innovation have failed, geometric morphometrics offers a wholly new, mainly geometric version of this discourse that we find extremely promising.

4.1 *Symmetry and asymmetry*

By this we mean *bilateral* symmetry, the left-right symmetry that (mostly) characterizes the vertebrates we study (along with the vertebrates we teach, or treat in hospitals, or write for). It is only an approximate symmetry, and the approximation is of considerable interest both for what varies and for what is constant. In the classic approaches, these studies proceed by measuring the same things twice: left and right lengths, or angles, or areas, which are then contrasted by mere subtraction. Or the biologist may label the locus of unpaired structures, the “midline”, and then contrast left and right displacements from this midline.

GMM improves on this unimpressive lexicon in at least four ways. First, it provides the language for describing *all* of the ways that the form differs from its own mirror-image at once: not only the difference in sizes left vs. right, but also the differences in their shapes, and further the differences in their positions with respect to the midline, both orthogonal to it and along it, and finally the nonflatness (bending) of that midline itself. Second, it provides statistics for testing any or all of these against the competing hypothesis of pure hazard, the reassurance (or not) that the apparent dissymmetry is not just a matter of wish-fulfillment on the investigator’s part, an interpretation of noise. Third, it allows us to fuse the scrutiny of asymmetry with the conceptual tools of integration to be reviewed presently. Where is the symmetry of the form more tightly regulated, where less so? (Why does the normal human pair of eyes differ so much more in lateral relationship than in vertical relationship?) Where dissymmetry is itself variable, in what dimensions it is most variable? When symmetry is “directional”, right larger than left or otherwise systematically different in form, how is this directionality expressed, and what is conserved? Fourth, as we shall see, the biologist is not the only organism interested in symmetry; it is also a central construct in the study of animal perception (as in, for instance, mate choice studies). How does the perceiving organism perceive asymmetry, and what are the similarities and dissimilarities between the conspecific contemplation of asymmetry and its scientific study? In all of these contexts, asymmetry is interpreted

in close analogy to its use in standard biometrics (Palmer and Strobeck 2003)—asymmetry in GMM is the shape difference between a form and its mirror image (Mardia *et al* 2000); directional asymmetry (DA) is the average of this, and fluctuating asymmetry (FA), the principal correlate of developmental stress, is the deviation of any case's asymmetry from DA.

4.2 Integration

The topic of morphological integration exploded into biometry only in 1958 (Olson and Miller 1958), a surprisingly recent date. That introduction came, likewise surprisingly, as an adaptation of techniques from the psychological sciences, specifically, factor analysis, with its emphasis on the essential interchangeability of variables (in contrast to individuality at the level of specimens). From its earliest origins in the 1970's, geometric morphometrics aimed to redress this imbalance, to put an attention to *variables* back at the core of the enterprise. There has resulted a rich language for the specification of *how the form is integrated*.

We have, for instance, a language of the spatial *scale* of features. Predictable changes (plastic or otherwise) can be uniform over the entire organism, or concentrated in accordance with any regionalization. Integration can be more intense within organs than between—the phenomenon called *modularity* (e.g. Needham 1933; Riedl 1978; Schlosser and Wagner 2004); GMM has powerful tools for exploring claims of that sort (e.g. Mitteroecker and Bookstein 2007, 2008). Integration is modified by a wide range of potentially or actually perturbing biological processes—growth, ecophenotypy, sexual dimorphism—any of which can disrupt stable patterns of integration (Mitteroecker and Bookstein 2009) or else leave them alone in the course of inculcating other, different patterns. (Think of “family resemblances” in genealogy, patterns like the specifically genetic prognathism in European royalty known as the Habsburg lip: integration along the biographical axis, the pedigree that is completely oblique to all the other developmental processes simultaneously operating.)

Indeed, in its current (2008) incarnation, the GMM concept of integration may be powerful enough to stand its ground in a new fundamental antinomy of evo-devo studies. In analysis of *integration versus plasticity*, GMM can contribute strength on both sides. Integration is the assertion of pattern *out of a grand mean form*, a form without memory, while plasticity, in the sense we are using it here, is the complementary statistic: a relentless influence of immediate past on future history. Technically, models for integration remain models of variation around a mean form; the distinction we are attempting to make is between those models and models of change from the immediately preceding value, whereas integrated systems “regress to the mean”.

5. Insights from GMM in studies of the plasticity of the human face

Under this heading, we would like to introduce some of our previous studies on human facial shape examining processes and depicting patterns attributable to sex and stress, and to their perception.

5.1 Asymmetry

We start with an example on the causes and continue with one on the consequences of asymmetry. We applied the Procrustes method for asymmetry introduced above (Mardia *et al* 2000) in an elegant natural experiment (Schaefer *et al* 2006b) studying the behaviour of a tightly regulated part of the cranium (the dentition): In a Croatian Adriatic island population (island of Hvar) we found that fluctuating dental arch asymmetry was higher than in a reference population from the Croatian capital (Zagreb), and increased within the island population with the extent of inbreeding. The study design allowed an examination of the influence on asymmetry of both environmental circumstances (by keeping the breeding factor constant and varying the environment with two exogamous samples from Hvar and from Zagreb) as well as of inbreeding (by keeping the environment—Hvar—constant and varying the extent of inbreeding). The environmental impact on FA seemed to exceed the inbreeding effect inasmuch as we find considerable differences in magnitude of FA between the Zagreb and outbred Hvar group in both jaws, but a weaker signal for the “within-Hvar comparisons”, especially in the upper jaw. The latter may be due to a stronger morphological integration of the upper jaw into the craniofacial complex and thus a higher sensitivity to developmental perturbations in the lower jaw. In keeping with the “widely held—yet poorly substantiated—belief that FA can act as a universal measure of developmental stability and predictor of stress” (Lens *et al* 2002) we assumed that FA signalled developmental precision. Yet, our samples yielded consistent differences in the extent of directionality depending on environmental and genetic stress levels. In our data, directional asymmetry appears with environmental stress in the lower jaw and is present in both jaws with further inbreeding stress, and FA increases with environmental stress in both jaws and with additional inbreeding stress in the lower jaw. These results do not replicate the transition from FA to directional asymmetry with increasing stress, but rather demonstrate significant directional asymmetry to co-occur with it, and therefore support the notion that DA itself may be an indicator of stress (Graham *et al* 1993).

As an example of the *consequences of asymmetry*, we found a negative association between FA and facial and body attractiveness (Schaefer *et al* 2006a)—not surprising since

perceived attractiveness in humans is proposed to be evidence of developmental and possibly hormonal health as well (e.g. Symons 1995; Thornhill and Møller 1997; Thornhill and Grammer 1999). The idea is that signals will evolve to be honest because selection for mate choosers using true fitness indicators in mate choice generate selection for competitive displays that signal honestly (e.g. Zahavi 1975; Zahavi and Zahavi 1997). Thus, mate choosers are expected to be selected to pay attention mainly to physical traits that honestly, or uncheatably, advertise mate value. In this sense, physical attractiveness could be considered as an honest Zahavian signal of phenotypic and genetic quality. Likewise intriguing is a finding that we were able to determine using a shape regression—a linear regression function calculated for every shape coordinate separately, whereby the slopes of the functions predict the shape change that occurs within one unit of that independent variable. Our data revealed that shape changes in the face which are associated with higher attractiveness ratings result in a configuration that appears generally rounder, with a more gracile nose and chin and larger lips (figure 1)—characteristics which are proposed to be facial estrogen markers (Johnston and Franklin 1993; Symons 1995).

Likewise, shape changes that relate to decreasing perceived female attractiveness resemble in many aspects features linked with the adult male face. Specifically, an

overall elongation, relatively smaller eyes, a larger nose, and a more pronounced region of mandible are some of the classic sexually dimorphic characteristics in men, linked to extended facial growth (hypermorphosis) and other developmental processes under the influence of testosterone (Enlow 1996; Rosas and Bastir 2002; Schaefer *et al* 2004; Bulygina *et al* 2006). Such features present in female faces might therefore indicate an estrogen to testosterone ratio shifted towards higher testosterone, and serve to trigger the less favourable physical evaluation we have found in the men's perception (Schaefer *et al* 2006a).

Along the same line lies also the next study we are going to report in the course of this paper, our attempt to disentangle some of the complexities of craniofacial sexual dimorphism (Schaefer *et al* 2004). Craniofacial dimorphism in primates has been documented by many authors over the years. There is some variation in patterns of dimorphism between species, and a broad association between body mass dimorphism and the overall magnitude of dimorphism among craniofacial dimensions. Apart from these generalizations, neither the degree to which patterns of craniofacial dimorphism are correlated with taxonomy, nor the variation in patterns within and among craniofacial regions, is well understood or agreed on (Plavcan 2002). We knew that the extent of sexual dimorphism in size has associations with the mating system (Plavcan and van

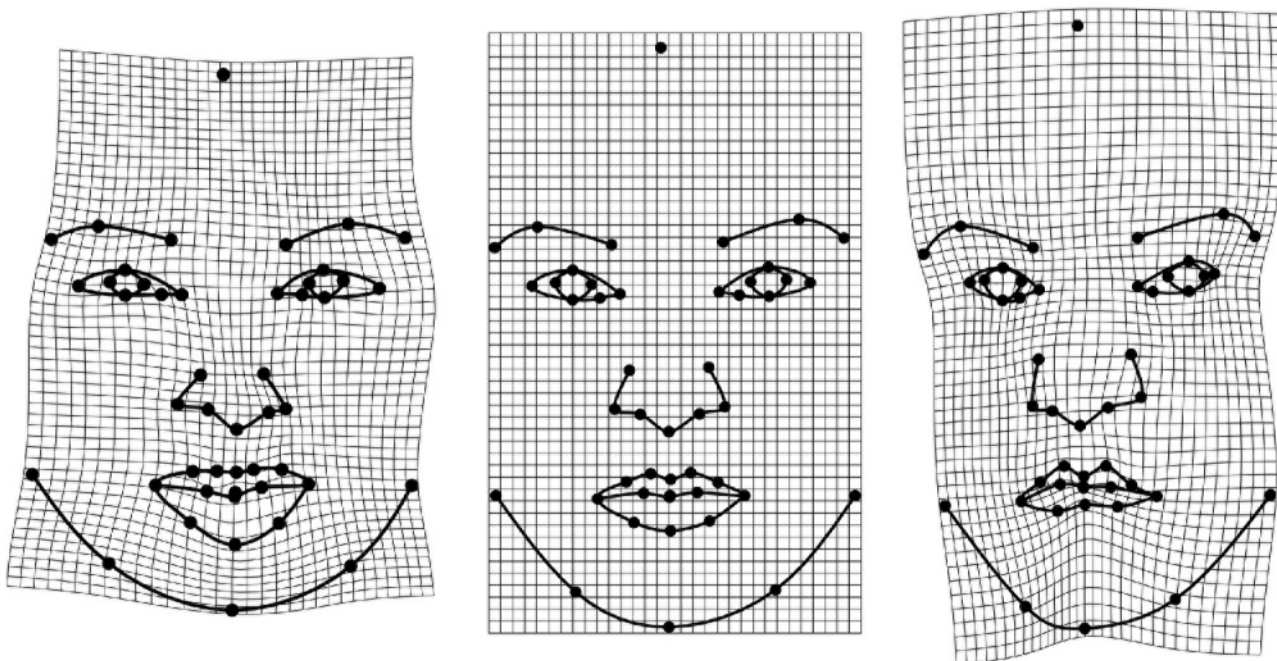


Figure 1. Visualization of the shape regression on perceived female facial attractiveness by thin-plate spline deformation grids. The middle panel, with an undeformed grid, is the average landmark configuration (consensus). Neighbouring panels show the predicted transformation in both directions from this consensus. The deformations correspond to an increase (left) or a decrease (right) by ten units on the attractiveness scale. (The factor of ten is chosen for readability of the warp; the actual sample variation of ± 2 units is too small to depict the distortions legibly.) Figure modified according to Schaefer *et al* (2006a).

Schaik 1997), and we assumed that the shape difference does not arise only from ontogenetic scaling (Shea 1983, 1986). For *Gorilla* and *Pongo*, the social system primarily involves high intensity male-male competition (Plavcan and van Schaik 1997) and marked differences in weight and size dimorphism. This might be achieved optimally in the given amount of time by strictly continuing ontogenetic growth. We therefore expected growth allometry to fundamentally contribute to their sexual dimorphism. Conversely, smaller size dimorphism (in *Pan* and *Homo*) is associated with lower-intensity male-male competition, and the shape differences between the sexes might be less allometric. We assessed the presence of ontogenetic scaling in hominoid crania using the angle formed between the vector of regression coefficients on the log of centroid size (the trajectory of ontogenetic allometry for all individuals within each species) and the vector between the mean adult male and female for each species. The angle ranged from very low (7.5°), to moderate (27.0°), to high (40.3°), to very high (62.1°) in *P. pygmaeus*, *H. sapiens*, *P. troglodytes*, and *P. paniscus*, respectively. We concluded from this that ontogenetic scaling contributes more to sexual dimorphism in *Pongo* than in either of the species of *Pan*, with humans mediating. Cobb and O'Higgins (2007) have just extended these results by noting that the difference emerges only subsequent to the eruption of the second molars. We ruled out a functional mechanical constraint as a likely explanation, nor were these morphological deviations from allometry likely to be miscellaneous developmental by-products in the course

of relatively slower growth. We suspected, rather, that they operate in sexual selection. Since the smaller magnitude in (body) size dimorphism is related to low-intensity male-male competition (in comparison to high-intensity male-male competition in gorilla and orangutan Plavcan and van Schaik 1997), more subtle strategies of sexual selection, such as sperm competition and female choice might play a role. We expected the high contribution of the non-allometric components to sexual dimorphism in *Pan* and *Homo* to be associated with male sexual attractiveness, such as enlarged cheekbones, 'testosterone markers', which are linked to sexual attractiveness.

One of the main shape differences associated with the late postnatal divergence of the male and female ontogenetic shape trajectories identified by Cobb and O'Higgins (2007) was an increase in the width of the zygomatic region associated with attachment of the masseter muscle and the underlying temporalis muscle. Similar dimorphisms have previously been reported in *Pan* (Weston *et al* 2004), and *Pongo* (Masterson and Leutenegger 1992; Hens 2003, 2005), as arising from divergent sex trajectories. Recently (Schaefer *et al* 2007) we have specifically scrutinized the craniofacial shape difference between a simulated purely allometrically grown member of the taxon and the actual mean male configuration by projecting the mean male onto the species-specific axis of allometry (figure 2).

In this way, we were able to show the impact of the non-allometric component on the species specific craniofacial shape of a male, and we demonstrated that the non-allometric

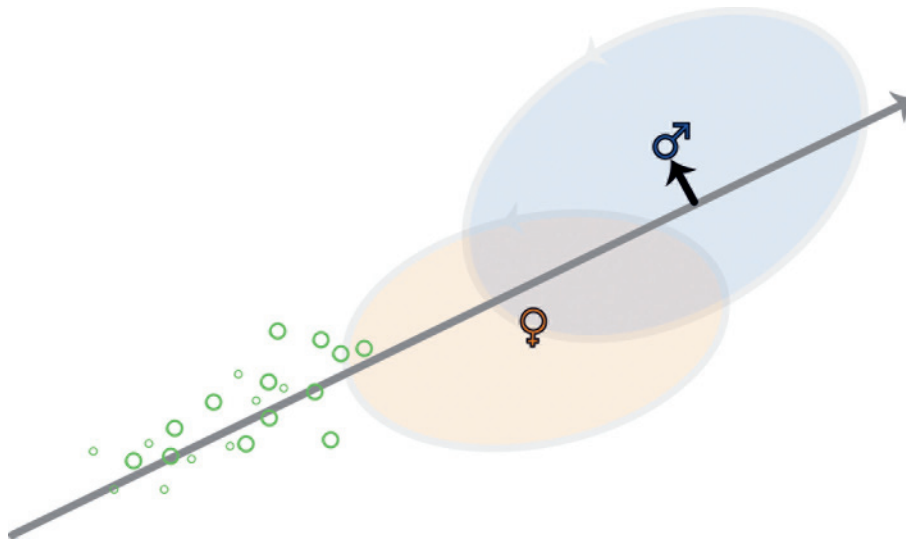


Figure 2. Scheme for the computation of “pure non-allometry” in males based on subadult, female and male cranial specimens (represented by more than 300 landmarks each) of one species plotted along the first two PCs in Form-space. The projection (short black arrow) of the male consensus configuration (♂ symbol) onto the species-specific axis of ontogenetic allometry (grey diagonal arrow) gives the shape differences between an artificially purely allometrically grown member of the species and the actual mean male. Small circles: subadult specimens; ♀, mean female; ♂, mean male; ellipses represent sample scatters.

sexual dimorphism component in *Homo* and chimpanzee males operates mainly in features that pertain to mid-facial width (Schaefer *et al* 2007). This is somewhat in accordance with a recent study by Weston and colleagues (2007) who found that at puberty, the height of the upper face, between the upper dental arch and the supraorbital ridge, develops differently in males and females, and that these differences are not explicable in terms of sex differences in body size, but suggest that adult men have relatively shorter upper faces for their breadth compared to women. Like Weston, we surmise that these observations have important implications for the role of sexual selection in the evolution of anthropoid faces and for theories of human facial attractiveness, namely that it was male competition driving the allometric component (sexual dimorphism in size), and female choice the non-allometric one (sexual dimorphism in shape). Clearly, the necessary next step would be to test this with living humans, or their facial photographs. We would expect to find the same pattern in sexual dimorphism as in the osteological specimens. Moreover, we predict male attractiveness ratings to pertain to shape features going not so much with larger body size (allometric scaling) but rather with the non-allometric ones. Only in the presence of such evidence would the notion of intersexual selection being responsible for the characteristics of this component gain empirical support.

6. Conclusions

We set out to briefly describe a methodological toolkit, criticize some of its assumptions as they related to the declared theme of this issue of the *Journal of Biosciences*, and sketch some of our own studies as they may relate to these same themes. Here at the end of the third component is a good place from which to return to the more abstract themes of the first two, the toolkit and its axioms.

We spoke, for instance, of two kinds of causation, one genetic, proximate (of inbreeding on dental asymmetry in Hvar), and one more ultimate and socio-environmental (aspects of sexual dimorphism that might be explicable by intersexual selection). Each of these is an emergent geometric trait, the consequence of decades of development over the course of which the claimed explanatory factors either might clearly still be operating (in the Hvar case) or might be so weak as to be quite debatable (the sexual selection case). Thus the meaning of causation is quite distinct between these two examples, and yet the biometrical method is essentially the same. It follows that there is some flexibility, or, as it were, plasticity in the way that biometric tools can be applied to biological questions.

In the case of sexual selection, what was originally plastic or otherwise subject to historical contingency must have been the specific selection of the non-allometric features

that are perceived to indicate competitive fitness. The data analysis begins from a “frozen” state in which some early adaptive radiation of varieties of sexual selection must have been resolved, possibly by accident, in the direction of shape effects that we just reviewed. The vector of non-allometric dimorphism in figure 2 cannot itself be argued to indicate any specific excellence of function. Rather, it is what it is, without any contemporary functional justification, so that whatever mechanisms preserve it must themselves be frozen into developmental programmes. (We exclude the realm of cultural determinism, as by advertising, which is characterized between reversals by an explanatory time-scale many orders of magnitude too short for the processes we are imagining to be operating.) The effectiveness of the morphometric (which is to say, biometric) method for an analysis in this spirit thus confirms the applicability of GMM to small changes at the end-stage of any process, integrated or plastic, as long as the remaining variation appears to be centered, subject to regression to the mean across generations, and so on.

By contrast, the analysis of dental asymmetry on Hvar need not be interpreted as frozen plasticity in this way. Dentofacial symmetry has clear implications for fitness: for instance, a hominid that can chew tough food with equal efficiency on both sides of the mouth is likely to be better-nourished than his asymmetric conspecifics, and thus to better protect his offspring so that they are likelier to survive to reproductive maturity in turn. The Hvar findings can be read as implying that the dental arches are integrated, not plastic: that displacement from an arguably optimal condition requires a perturbation of cause (here, the inbreeding, the absence of buffering).

So the single methodology of GMM is, in our reading, quite catholic as regards this evo-devo divide between integration and plasticity as highest-level explanations, as mantras for the generation of follow-on hypotheses or imagined mechanisms explaining the empirical findings by recourse to supplementary questions. This evenhandedness is, in our view, a very great merit of GMM, not a drawback. In a field whose theoretical basis is as fraught as that of today’s evo-devo sciences, studies to weigh the two great competing styles of explanation will have to rely on a methodology that is neutral as regards the better candidate. GMM seems to be such a methodology, neutral as to the difference between integration and plasticity as fundamental processes and thus capable of weighing their separable contributions to a net explanation. (Recall how in the case of sexual dimorphism, there was evidence of both types of process. Allometry is integrative, is in fact the “holotype” of an integrative argument; but non-allometric dimorphism is surely plastic, historically conditioned, *over exactly the same space of shape coordinates*.) Size, the causal factor driving allometry, is indeed sensitive to environment,

which makes it “plastic” in the conventional sense but not in the sense of this paper, as mammalian growth, at least, is typically set early in life.

One can interpret this flexibility of the method as a testimonial to the robust strengths of the biometrical point of view, or as a deeper declaration about the interdisciplinary relevance of the statistical sciences across the world of quantitative natural sciences more generally. What we are seeing here is surely not the quantitative laws of physics—this is a very old point, made by Karl Pearson in 1911 already in his *Grammar of Science*—but a different sort of fundamental rhetoric, the value of concentrating on the variation, not the ideal. Aristotle, not Plato. To the extent that these quantifications lead to a consensus about progress, it must be that the distinction between integration and plasticity as explanations is compatible with the idea of an empirical quantitative biology. This is possibly the most optimistic possible deduction of all, and a splendid outcome of the conference out of which the papers in this publication arose.

Acknowledgements

We would like to express our gratitude to the editors and organizers of the Estuary Island Discussion Meeting on Phenotypic and Developmental Plasticity, Vidyanand Nanjundiah and Stuart Newman, for having inspired and enabled this piece of work, as well as to Philipp Mitteroecker and Philipp Gunz for their contribution to all the previous work mentioned in the course of this paper. We also thank two reviewers for helpful comments on an earlier version of this manuscript. Support by EU FP6 Marie Curie Actions (EVAN) MRTN-CT-2005-019564 and by the University of Vienna.

References

- Bookstein F L 1991 *Morphometric tools for landmark data: geometry and biology* (Cambridge: Cambridge University Press)
- Bookstein F L 1996 Biometrics, biomathematics and the morphometric synthesis; *Bull. Math. Biol.* **58** 313–365
- Bookstein F L 2009 Geometric morphometrics for virtual anthropology; in *virtual anthropology: A guide to a new interdisciplinary field* (eds) G W Weber and F Bookstein (Wien, New York: Springer-Verlag) capital 5 (in press)
- Bulygina E, Mitteroecker P and Aiello L C 2006 Ontogeny of facial dimorphism and patterns of individual development within one human population; *Am. J. Phys. Anthropol.* **131** 432–443
- Cobb S N and O’Higgins P 2007 The ontogeny of sexual dimorphism in the facial skeleton of the African apes; *J. Hum. Evol.* **53** 176–190
- Dryden I L and Mardia K V 1993 Multivariate shape analysis; *Sankya* **55** 460–480
- Dryden I L and Mardia K V 1998 *Statistical shape analysis* (Chichester: Wiley)
- Elsasser W 1998 *Reflections on a theory of organisms* (Baltimore: The Johns Hopkins University Press)
- Enlow D H 1996 *Essential of facial growth* (Philadelphia: W.B. Saunders Company)
- Gaddis J L 2004 *The landscape of history: How historians map the past* (New York: Oxford University Press)
- Goodall C R 1991 Procrustes methods and the statistical analysis of shape; *J. R. Stat. Soc. B.* **53** 285–340
- Gower J C. 1966 Some distance properties of latent root and vector methods used in multivariate analysis; *Biometrika* **53** 325–338
- Gower J C 1975 Generalized Procrustes analysis; *Psychometrika* **40** 33–50
- Graham J H, Roe K E and West T B 1993 Effects of lead and benzene on the developmental stability of *Drosophila melanogaster*; *Ecotoxicology* **2** 185–195
- Hallgrímsson B, Willmore K and Hall B K 2002 Canalization, developmental stability, and morphological integration in the primate limbs; *Yearb. Phys. Anthropol.* **45** 131–158
- Hens S N 2003 Growth and sexual dimorphism in orangutan crania: a three-dimensional approach; *Am. J. Phys. Anthropol.* **121** 19–29
- Hens S N 2005 Ontogeny of craniofacial sexual dimorphism in the orangutan (*Pongo pygmaeus*) I: face and palate; *Am. J. Primatol.* **65** 149–166
- Johnston V S and Franklin M 1993 Is beauty in the eye of the beholder? *Ethol. Sociobiol.* **14** 183–199
- Kent J T 1994 The complex Bingham distribution and shape analysis; *J. R. Stat. Soc. B.* **56** 285–299
- Lens L, Van Dongen S, Kark S and Matthyssen E 2002 Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies?; *Biol. Rev.* **77** 27–38
- Ludwig W 1932 *Das Rechts-Links Problem im Tierreich und beim Menschen* (Berlin: Springer)
- Marcus L F, Corti M, Loy A, Naylor G J P and Slice D 1996. *Advances in morphometrics; NATO ASI Series* (New York: Plenum Press)
- Mardia K V, Bookstein F L and Moreton I J 2000 Statistical assessment of bilateral symmetry of shapes; *Biometrika* **87** 285–300
- Masterson TJ and Leutenegger W 1992 Ontogenetic patterns of sexual dimorphism in the cranium of Bornean orang-utans (*Pongo pygmaeus pygmaeus*); *J. Hum. Evol.* **23** 3–26
- Mitteroecker P and Bookstein F L 2007 The conceptual and statistical relationship between modularity and morphological integration; *Syst. Biol.* **56** 818–836
- Mitteroecker P and Bookstein F L 2008 The evolutionary role of modularity and integration in the hominoid cranium; *Evolution* **62** 943–958
- Mitteroecker P and Bookstein F L 2009 The ontogenetic trajectory of the phenotypic covariance matrix, with examples from craniofacial shape in rats and humans; *Evolution* (in press)
- Needham J 1933 On the dissociability of the fundamental processes in ontogenesis; *Biol. Rev.* **8** 180–223
- O’Higgins P 2000a Quantitative approaches to the study of craniofacial growth and evolution: advances in morphometric techniques; in *Vertebrate ontogeny and phylogeny: implications for the study of hominid skeletal evolution* (eds) P O’Higgins and M Cohn (London: Academic Press) pp 163–185

- O'Higgins P 2000b The study of morphological variation in the hominid fossil record: biology, landmarks and geometry; *J. Anat.* **197** 103–120
- O'Higgins P and Jones N 1998 Facial growth in *Cercocebus torquatus*: an application of three dimensional geometric morphometric techniques to the study of morphological variation; *J. Anat.* **193** 251–272
- Olson E C and Miller R L 1958 *Morphological integration* (Chicago: University of Chicago Press)
- Palmer A R and Strobeck C 2003 Fluctuating asymmetry analysis revisited; in *Developmental instability: causes and consequences* (ed.) M Polak (New York: Oxford University Press) pp 279–319
- Phenotypic plasticity 2008 *Encyclopædia Britannica*. Retrieved December 09, 2008, from *Encyclopædia Britannica Online*: <http://www.britannica.com/EBchecked/topic/455640/phenotypic-plasticity>
- Pigliucci M and Preston K 2004 *Phenotypic integration: Studying the ecology and evolution of complex phenotypes* (New York: Oxford University Press)
- Plavcan J M 2002 Taxonomic variation in the patterns of craniofacial dimorphism in primates; *J. Hum. Evol.* **42** 579–608
- Plavcan J M and van Schaik C P 1997 Intrasexual competition and body weight dimorphism in anthropoid primates; *Am. J. Phys. Anthropol.* **103** 37–68
- Polak M 2003 *Developmental instability: causes and consequences* (New York: Oxford University Press)
- Riedl R J 1978 *Order in living organisms* (New York: John Wiley)
- Rohlf F J 2000a On the use of shape spaces to compare morphometric methods; *Hystrix Italian J. Mammol.* **11** 9–25
- Rohlf F J 2000b Statistical power comparisons among alternative morphometric methods; *Am. J. Phys. Anthropol.* **111** 463–478
- Rohlf F J 2003 Bias and error in estimates of mean shape in geometric morphometrics; *J. Hum. Evol.* **44** 665–683
- Rohlf F J and Bookstein F L (eds) 1990 *Proceedings of the Michigan morphometrics workshop. Special Publication No. 2* (Ann Arbor: The University of Michigan Museum of Zoology)
- Rohlf F J and Slice D E 1990 Extensions of the Procrustes method for the optimal superimposition of landmarks; *Syst. Zool.* **39** 40–59
- Rosas A and Bastir M 2002 Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex; *Am. J. Phys. Anthropol.* **117** 236–245
- Schaefer K, Grammer K, Fink B, Mitteroecker P, Gunz P and Bookstein F L 2006a Female appearance: facial and bodily attractiveness as shape; *Psychol. Sci.* **48** 178–205
- Schaefer K, Lauc T, Mitteroecker P, Gunz P and Bookstein F L 2006b Dental arch asymmetry in an isolated community; *Am. J. Phys. Anthropol.* **129** 132–142
- Schaefer K, Mitteroecker P, Gunz P, Bernhard M and Bookstein F L 2004 Craniofacial sexual dimorphism patterns and allometry among extant hominoids; *Ann. Anat.* **186** 471–478
- Schaefer K, Mitteroecker P, Gunz P, Bernhard M and Bookstein F L 2007 “Maleness” reconsidered: hominoid craniofacial sexual dimorphism; *Am. J. Phys. Anthropol.* **132** S44 206–207
- Schlusser G and Wagner G P 2004 *Modularity in development and evolution* (Chicago: The University of Chicago Press)
- Shea B T 1983 Allometry and heterochrony in the African apes; *Am. J. Phys. Anthropol.* **62** 275–289
- Shea B T 1986 Ontogenetic approaches to sexual dimorphism in anthropoids; *J. Hum. Evol.* **1** 97–110
- Slice D E (ed.) 2005 *Modern morphometrics in physical anthropology* (New York: Kluwer Academic Publishers)
- Slice D E 2007 Geometric morphometrics; *Annu. Rev. Anthropol.* **36** 261–281
- Smith L I 2002 http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. Accessed Dec. 7, 2008
- Symons D 1995 Beauty is in the adaptations of the beholder: the evolutionary psychology of human female sexual attractiveness; in *Sexual nature, sexual culture* (eds) P R Abramson and S D Pinkerton (Chicago, London: The University of Chicago Press) pp 80–120
- Thompson D'A W 1917 *On growth and form* (London: Macmillan)
- Thornhill R and Grammer K 1999 The body and face of a women: one ornament that signals quality? *Evol. Hum. Behav.* **20** 105–120
- Thornhill R and Møller A P 1997 Developmental stability, disease and medicine; *Biol. Rev.* **72** 497–548
- Weston E M, Friday A E, Johnstone R A and Schrenk F 2004 Wide faces or large canines? The attractive versus the aggressive primate; *Proc. R. Soc. London B* **271** 416–419
- Weston E M, Friday A E, Liò P 2007 Biometric evidence that sexual selection has shaped the hominin face; *PLoS ONE* **2** e710
- Weyl H 1952 *Symmetry* (Princeton: Princeton University Press)
- Wikipedia 2008 http://en.wikipedia.org/wiki/Principal_components_analysis. Accessed Dec. 7, 2008
- Zahavi A 1975 Mate selection—selection for a handicap; *Theor. Biol.* **53** 205–214
- Zahavi A and Zahavi A 1997 *The handicap principle* (New York: Oxford University Press)

ePublication: 13 February 2009