



How Exactly Did the Nose Get That Long? A Critical Rethinking of the Pinocchio Effect and How Shape Changes Relate to Landmarks

Christian Peter Klingenberg¹

Received: 9 July 2020 / Accepted: 22 October 2020 / Published online: 30 October 2020
© The Author(s) 2020

Abstract

The Pinocchio effect has long been discussed in the literature on geometric morphometrics. It denotes the observation that Procrustes superimposition tends to distribute shape changes over many landmarks, even though a different superimposition may exist for the same landmark configurations that concentrates changes in just one or a few landmarks. This is widely seen as a flaw of Procrustes methods. Visualizations illustrating the Pinocchio effect use a comparison of the same pair of shapes superimposed in two different ways: in a resistant-fit superimposition that concentrates the shape difference in just one or a few landmarks, and in Procrustes superimposition, which distributes differences over most or all landmarks. Because these superimpositions differ only in the non-shape aspects of size, position and orientation, they are equivalent from the perspective of shape analysis. Simulation studies of the Pinocchio effect usually generate data, either single pairs or larger samples of landmark configurations, in a particular superimposition so that differences occur mostly or exclusively at just one or a few landmarks, but no steps are taken to remove variation from size, position and orientation. When these configurations are then compared with Procrustes-superimposed data, differences appear and are attributed to the Pinocchio effect. Overall, it is ironic that all manifestations of the Pinocchio effect in one way or another rely on differences in the non-shape properties of position and orientation. Rigorous thinking about shape variation and careful choice of visualization methods can prevent confusion over this issue.

Keywords Geometric morphometrics · Landmarks · Pinocchio effect · Procrustes superimposition · Shape

Introduction

Superimposition methods have an important role in geometric morphometrics for quantifying the differences between pairs of shapes or for aligning landmark configurations to the respective shape (tangent) spaces (Klingenberg 2020). Because shape is defined as all the geometric features in a configuration of landmarks except for its size, position, and orientation, superimposition methods are used to find a standardization of size, position, and orientation that minimizes some criterion quantifying the differences in the positions of landmarks relative to each other. Discrepancies in the positions of landmarks that remain after this standardization step can be ascribed to shape differences between the

respective landmark configurations. The most widely used method is Procrustes superimposition, where the criterion being minimized is the sum of squared distances between all the landmarks of the configuration (Boas 1905; Sneath 1967; Rohlf and Slice 1990; Goodall 1991; Dryden and Mardia 2016). This approach is at the core of geometric morphometrics and defines Procrustes distance, the metric used for quantifying differences between shapes, and also determines Kendall's shape space, which forms the theoretical foundation for much of statistical shape analysis (Kendall 1984; Bookstein 1996; Kendall et al. 1999; Dryden and Mardia 2016; Klingenberg 2016, 2020).

A topic that periodically has surfaced in the literature on geometric morphometrics is a critique of the Procrustes superimposition based on a widespread perception that it tends to distribute shape changes erroneously over many or all landmarks (Siegel and Benson 1982; Chapman 1990; Rohlf and Slice 1990; Klingenberg and McIntyre 1998; Walker 2000; von Cramon-Taubadel et al. 2007; Zelditch et al. 2012; Hallgrímsson et al. 2015; Tatsuta et al. 2018;

✉ Christian Peter Klingenberg
cpk@manchester.ac.uk

¹ School of Biological Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK

Palci and Lee 2019; Thulman 2019; Courtenay et al. 2020; Fox et al. 2020). Because this is often illustrated with the example of Pinocchio's nose growing longer after he told a lie, this critique has often been discussed under the heading of the "Pinocchio effect" (Chapman 1990; Klingenberg and McIntyre 1998; Walker 2000; Slice 2005; Gill et al. 2007; von Cramon-Taubadel et al. 2007; Zelditch et al. 2012; Hallgrímsson et al. 2015; Fruciano 2016; Tatsuta et al. 2018; Palci and Lee 2019; Thulman 2019; Courtenay et al. 2020; Fox et al. 2020).

This paper investigates the logic that underlies the thinking about the Pinocchio effect and some consequences of that logic. As it turns out, such examination of common ways of thinking, both about shape analysis and about Pinocchio's nose, can yield some rather intriguing insights. When it comes to understanding geometric morphometrics, the role of the Procrustes superimposition, as well as the Pinocchio effect, the key question is how shape changes do or do not relate to specific landmarks. A point of particular importance is the visualization of shape changes, which profoundly influences how morphometricians think about shape variation (Klingenberg 2013b). Thought experiments and back-of-the-envelope simulations can easily produce misconceptions if they are based on misguided visualizations. Likewise, numerical simulation studies can provide misleading results, depending on how models are set up and results are interpreted.

Shape, Icons, and Shape Variation

Shape is defined as all the geometric features of an object except for its size, position, and orientation (Goodall 1991; Dryden and Mardia 2016). Formally, shape can be represented by an equivalence class: the set of objects that differ only in their sizes, positions and orientations (Goodall 1991; Dryden and Mardia 2016). Shape encompasses all those geometric properties shared by all the members of that set. For objects represented as configurations of landmarks, these properties include, for example, the proportions among distances between pairs of landmarks, angles formed by lines drawn between landmarks, and the relative arrangements of portions of the landmark configuration representing different parts of the biological structure under study. Shape is inherently an abstract property, as we can only see the object or its representation as a particular instance from the equivalence class with a particular size and in a particular location and orientation. These can be the physical size, location and orientation of the actual object itself or those of a representation on a computer screen, on paper, or as a 3D model or hologram—each of those representations has its own specific size, position and orientation. Such an instance of a shape, with a particular size, position and

orientation, can be denoted by the term "icon" (Dryden and Mardia 2016). For example, a particular fly wing and its image on a computer screen are two different icons for the shape of that fly wing.

Much of our thinking about shape and shape variation actually involves icons rather than shapes, and even statistical analyses of shape variation use particular icons to represent shapes. The key point, then, is to make sure that the conclusions drawn from those analyses do not depend on the inherently arbitrary choice of particular icons.

Visualizations of shape changes usually involve a pair of icons, one each from the equivalence classes of the starting and the target shape. How the two icons of the pair relate to each other, apart from the shape change, is completely open, and can affect the intuitive appearance of the visualization substantially (Klingenberg 2013b). The two icons may be shown on top of each other in the superimposition resulting from a Procrustes fit, in which the effects of the shape change tend to be distributed across most or all of the landmarks (Fig. 1b). An alternative is to show the two icons in an alignment that concentrates the effects of the shape change into as few landmarks as possible,

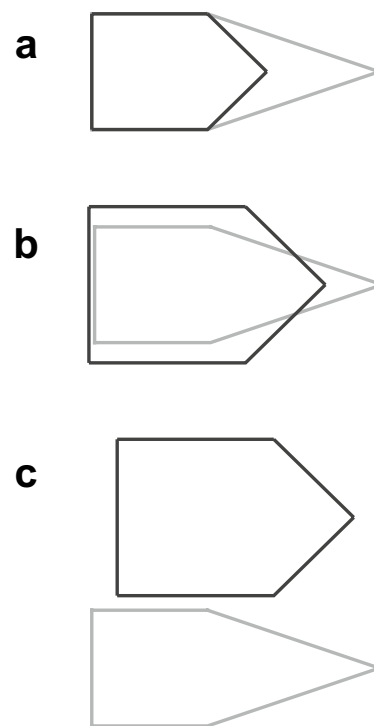


Fig. 1 Three visualizations of the same shape change. (modified from Klingenberg 2013b). All three involve the same shape change, but differ in how the two icons of the respective pair (grey and black lines) relate to each other. **a** The two icons are in the arrangement that concentrates all the differences in a single landmark. **b** The two icons are in the superimposition from a Procrustes fit, which distributes the differences over all the landmarks. **c** The two icons are shown side by side, without any overlap

according to some principle of parsimony (Fig. 1a). Or quite simply, the two icons may be shown side by side, so that they do not overlap (Fig. 1c). This latter option has the psychological advantage that it reduces the temptation for viewers to ascribe shape changes to individual landmarks (Klingenberg 2013b). Note, however, that there is no inherent reason whatsoever to prefer any one of those options over the others, or over the limitless other possibilities of displaying the pair of icons in different relationships to one another. From the standpoint of shape analysis, they are all equivalent because they involve the same pair of shapes.

Most morphometric analyses are based on shape data represented as particular icons, with some particular scaling, position and orientation chosen to be convenient for the analyses and visualization of results. In itself, this is not problematic. The key point is to keep in mind that those choices should not affect the results. Some outputs from statistical analyses of morphometric data can be interpreted as shape changes, and visualizing them again makes use of icons, whose sizes, positions and orientations can be chosen freely.

Collecting morphometric data means recording the coordinates of landmarks for the icons corresponding to the specimens under study. Some aspects of the information contained in each configuration of landmarks are relevant for morphometric studies, whereas others are not. The position and orientation of each configuration are clearly irrelevant from a biological perspective because the position just indicates where the specimen was relative to the camera or digitizing equipment and the orientation just says which way it was pointing during the data collection. Size, by contrast, is biologically relevant because it is an important organismal feature (Schmidt-Nielsen 1984) and may be of use in further analyses, for instance, to study allometry (Klingenberg 2016). An important task in morphometric studies is to separate the irrelevant aspects of variation from those that have biological importance and need to be retained for further analyses.

This task is usually achieved by using Procrustes superimposition (Boas 1905; Sneath 1967; Rohlf and Slice 1990; Goodall 1991; Dryden and Mardia 2016; Klingenberg 2020). It works by finding an alignment so that a landmark configuration fits as closely as possible to a target configuration. In this process, the first configuration is scaled, translated and rotated so that a measure of difference between the two configurations is minimized. The measure of difference is usually the sum of squared distances between corresponding landmarks. The differences remaining between landmark positions in the two configurations, which cannot be removed by changing scale position and orientation, can then be interpreted as manifestations of shape differences. An alternative explanation of Procrustes superimposition is that it is a procedure to quantify the shape distance between

a pair of landmark configurations or to fit a set of landmark configurations to Kendall's shape space (Klingenberg 2020).

Fitting landmark data to the shape space, or in practice the shape tangent space, imposes a number of constraints on the variation of landmark coordinates (Klingenberg 2020). These constraints have consequences for how shape changes appear in visualizations or in numerical representations of vectors in the shape tangent space. The constraints for translation always involve all landmarks simultaneously, because translation inevitably affects all landmarks simultaneously and equally. Usually, the constraints for scaling and rotation also affect all landmarks jointly. The only exception are landmarks that, for the shape at the tangent point, coincide precisely with the centroid of the landmark configuration, because that point is the center of rotation. These statements of constraints encapsulate, in a geometrically more rigorous form, the insight that displacements of landmarks are always relative to all the other landmarks in the configuration (Klingenberg 2013b).

The Pinocchio Effect: How It is Explained

The Pinocchio effect has been explained most concisely by Walker (2000) as the situation “when all of the shape variation occurs at a single landmark” (p. 686 f.), but it is usually presented visually using some hypothetical example involving a highly localized shape change. Chapman's (1990) original example compares two sea stars, one with all five arms intact and another one with one arm that was lost and has just begun to regenerate, and therefore is much shorter than the other four arms (the arms of both sea stars are in the same relative orientations). Walker (2000) used the ontogenetic shape change in the bird wrasse (*Gomphosus varius*) as an example, which is dominated by a disproportionate elongation of the snout. The prime example, of course, is Pinocchio's nose, which elongates drastically when he tells a lie. In this case, the comparison is between landmark configurations of Pinocchio's head when he is truthful and when he is lying. Other examples (Siegel and Benson 1982; Rohlf and Slice 1990; Palci and Lee 2019) use pairs of abstract configurations of landmarks with differences concentrated in one particular part, such as the simple example in Fig. 1. For any of those pairs of shapes, two visualizations of the respective shape difference are presented: one according to the Procrustes superimposition of one landmark configuration onto the other, and a second one where the majority of landmarks of the two configurations coincide and the shape change is concentrated in just one or a few landmarks (e.g. Figure 1a, b).

From comparing these visualizations of the difference between the pair of shapes, Chapman (1990) concluded that “... the least-squares algorithm distributed the lack of fit

among the landmarks, whereas the [resistant-fit algorithm] analysis represented the fit correctly” (p. 260 f.). From this contrast, it appears clearly that Chapman deemed it incorrect that the least-squares superimposition distributed differences across all landmarks. Similarly, Walker (2000) noted that the resistant-fit method “correctly locates most of the shape difference between a juvenile and initial phase bird wrasse at the snout, whereas a least-squares superimposition spreads this difference at the snout across many of the other landmarks” (p. 687) and further mentioned that “a least-squares superimposition can mistakenly identify the location of shape difference between a pair of figures” (p. 688). Likewise, as consequences of Procrustes superimpositions of two lateral views of Pinocchio’s head with short and long nose, Hallgrímsson et al. (2015) found “displacements of landmarks that were not actually displaced” (p. 572), “artifactual shape changes” (p. 572) and “artifactual variation” (p. 573). The choices of wording in these papers all imply clear judgements about which changes are correct and which ones are mistaken or artifactual (similar but less explicit statements can be found in many other papers). These judgements are presented without a detailed explanation of the logic upon which they are based—it appears that the authors saw the visualizations as self-evident and not needing further justification (and it seems that, for three decades, readers of the morphometrics literature have accepted them as such without questioning).

What exactly is the logic underlying those judgements? The visualizations each show a pair of shapes, represented as icons shown in two separate superimpositions. For one superimposition, the two icons are superimposed so that a maximum number of landmarks (in most cases, all but one) coincide precisely and all the change appears in the remaining ones (Fig. 1a), whereas for the other, the two icons are superimposed using least-squares Procrustes superimposition (e.g., Fig. 1b). Because both visualizations in each comparison include icons from the same pair of shapes, both visualizations of the shape difference in each comparison are equivalent. The difference between the two superimpositions therefore is only in the size, position and orientation of the two icons relative to each other. Far from one visualization of the shape difference being correct and the other artifactual, there is actually no basis in the theory of shape analysis for deeming either inherently better than the other.

The reason why a superimposition that concentrates the effects of a shape change in a single or a few landmarks (Fig. 1a) is so much more appealing for most viewers than the visualization of the same shape change that distributes its effects over most or all landmarks (Fig. 1b) seems to be that most people have a strong cognitive preference for ascribing the change to as few landmarks as possible. Such a parsimonious account of the observed changes appears more natural and understandable than a characterization of

change distributed over many landmarks. This preference for parsimonious interpretations of shape changes may also be connected to the real difficulty of most people, including experienced morphometricians, to fully grasp the idea that shape changes are not inherently linked to particular landmarks. That the allocation of shape changes to particular landmarks is a matter of parsimony has also been noted by Albert et al. (2003).

Palci and Lee (2019) presented a different variant of the critique of Procrustes superimposition invoking the Pinocchio effect. They present an example using two abstract configurations: rectangles with “spines” on the two long sides, constructed so that all the landmarks can coincide, except for the tip of one of the spines, which makes that spine bigger or smaller. They argue that the resistant-fit superimposition, which produces a fit where all but one landmarks are unchanged, corresponds to what most systematists would intuitively use as an alignment to identify the variable part of the structure and to characterize the difference between forms. Perhaps it is important to note that, for Palci and Lee (2019), the goal of comparing the forms is to define characters that can be used in cladistic analyses (an endeavor that has itself been questioned; Varón-González et al. 2020). Therefore, they can justify the preference for resistant-fit superimposition by invoking the preference for parsimonious representations of change that is fundamental to cladistic methods in general (e.g., Kitching et al. 1998). By contrast, the concept of shape does not figure prominently in their reasoning. This makes this type of study rather different from most morphometric studies.

How Do Shape Changes Relate to Landmarks?

When Pinocchio is lying, his nose lengthens dramatically, resulting in a change of the overall shape of his head and face from when he is honest. Whereas this premise is clear, it is much more difficult to assign this effect to specific landmarks. There is no way to decide objectively whether the landmark at the tip of the nose is moving forward and away from the landmarks on other parts of the face, or whether those landmarks together move posteriorly and thus away from the tip of the nose. Both descriptions characterize the shape change correctly, as each of them uses one icon for the starting shape (honest Pinocchio, with short nose) and one for the target shape (lying Pinocchio, with long nose). Because this shape change happens while Pinocchio may be moving around freely, there is no fixed coordinate system that could serve as a reference for the choice of a pair of icons. As a result, there are many pairs of icons that are equivalent in showing the same shape change, but that result in different displacements of the landmarks (e.g., Fig. 1).

Accordingly, there is no unique association between shape changes and landmark displacements.

The reason for this apparent ambiguity is that shape changes affect the arrangement of the landmarks relative to each other. It is therefore more helpful to consider shape changes to happen *between* the landmarks, rather than *at* particular landmarks. In a biological context, this means that shape changes result from changes in the tissues between the landmarks, not at the landmarks per se. Shape changes originate because the tissues surrounding the landmarks expand, contract, warp or distort so that they push or pull the landmarks in different directions. According to this view, the landmarks are dragged along passively by the processes that produce change in the interjacent tissues.

A particularly clear illustration of this idea are studies of leaf growth in which investigators draw or print a grid of points on the surface of a developing leaf and repeatedly take pictures as the leaf expands (Avery 1933; Saurer and Possingham 1970; Poethig and Sussex 1985; Granier and Tardieu 1998; Wang et al. 2011). Changes in the leaf shape and size can be analyzed based on the changes in the arrangement of the points in the grid, which is stretched and distorted to various degrees depending on the relative rates of growth in different directions at every location on the leaf. In this type of experiment, it is particularly evident that the points of the grid are moved passively by the growth processes in the expanding leaf, not by any active involvement of the markers themselves. Of course, the same processes that produce relative shifts among artificial markers also can displace landmarks that are defined anatomically. These landmarks are recognizable without experimental intervention, but their positions relative to each other (and in relation to the structure on which they are located) are also determined by variation in the tissues that surround the landmarks, not by the landmarks themselves.

The reasoning that shape changes originate from processes taking place in the tissues in which landmarks are situated, and that landmarks are displaced as a result, considers processes akin to morphogenetic movements, where tissues deform or move and therefore can push and pull around passive markers such as landmarks (Montell 2008; Kennaway et al. 2011). This raises the questions whether different types of developmental processes, such as patterning events where new boundaries or cell domains are specified in apparently homogeneous arrays of cells, might behave differently and constitute mechanisms for locally autonomous developmental processes. If so, such processes might produce variation in the resulting structures, and thus also landmarks, that is independent of surrounding tissues. Yet, the available information about these patterning mechanisms does not support this idea, because patterning inputs, such as morphogen gradients and other intercellular signaling processes, are themselves not acting strictly locally, but are

controlling patterning and growth throughout a region and are thereby integrating local patterns with the surrounding tissues (Sagner and Briscoe 2017). Also, patterning processes and morphogenetic movements interact (Heisenberg and Bellaïche 2013; Gilmour et al. 2017), further causing variation to have regional rather than strictly localized effects.

These considerations lead to the conclusion that, from a biological point of view, shape changes cannot be ascribed to individual landmarks but are associated inextricably with the space between them. Landmarks are not the agents of shape change, but changes in the surrounding tissues displace the landmarks and rearrange the configurations they form. At most, therefore, shape changes can be localized to a region between neighboring landmarks, but not to individual landmarks on either side of such a region.

Visualizations of Shape Changes

As mentioned above, visualizations of shape changes usually involve two icons representing the two shapes being compared. Even though it makes no difference for the shape comparison per se, the choice of icons from the equivalence classes corresponding to the two shapes can have a substantial effect on how a viewer perceives the shape change. Figure 1 shows three visualizations of the same shape change: the three black outlines represent icons from the equivalence class of one shape and the three gray outlines are icons from the equivalence class of another shape. Therefore, in principle, all pairwise comparisons between any of the black outlines and any of the gray outlines in the figure are equivalent because they show the same shape change (and there would be an infinity of other possible visualizations showing icons of the two shapes in different sizes, positions, and orientations). In practice, however, the three pairwise comparisons within the parts of the figure labeled (a), (b), and (c) are most relevant.

Discussions of the Pinocchio effect often contrast two visualizations: one that maximally concentrates differences at one or a few landmarks (Fig. 1a) and one that disperses the differences over most or all of the landmarks (Fig. 1b). Resistant-fit superimposition (Siegel and Benson 1982; Chapman 1990; Rohlf and Slice 1990) is a method for superimposing landmark configurations so that differences are concentrated in one or a few landmarks. By contrast, least-squares Procrustes superimposition (Goodall 1991; Dryden and Mardia 2016) tends to distribute changes over multiple landmarks because the least-squares criterion penalizes large differences in individual landmarks (there is a smaller overall penalty for many small differences). How much the visualizations obtained with different superimposition methods differ depends on the shape difference between landmark

configurations. Explanations of the Pinocchio effect often use landmark configurations for which there is a pair of icons that differ in the position of just one landmark, whereas all others coincide perfectly. As a result, there is an extreme difference between the visualization that concentrates the effects of shape variation in few landmarks, a single one in this case (Fig. 1a), and the visualization that distributes these effects over many landmarks (Fig. 1b).

That a visualization where the effects of a shape difference are concentrated in one or a few landmarks appeals more to viewers than one where changes are distributed over many landmarks seems to relate to a widespread preference for simple characterizations of shape changes. Such a preference, however, is entirely a matter of the perception of differences between pairs of icons by the human visual and cognitive systems, but it does not have a basis in the theory of shape analysis, for which all visualizations of a given shape change are equivalent.

There are visualization methods other than pairs of icons shown in some position on top of each other or side by side (Fig. 1). Especially transformation grids (Thompson 1961) using the thin-plate spline (Bookstein 1989) are relevant here, because they offer a way to avoid the visualization using a pair of icons (Fig. 2). Transformation grids use a different type of reasoning, because they involve an implicit comparison between a starting shape, for which the transformation grid is perfectly rectangular and all grid cells are usually square (Fig. 2a), and a target shape for which the grid is warped (Fig. 2b). This warping is the same, regardless

of the size, position, and orientation of the icons used for the starting and target shapes, as the grids take a size, position and orientation to match those of the corresponding icon. The warped transformation grid shows changes in the regions affected by shape changes, which may appear as relative expansions or contractions, shearing, bending and similar regional distortions. Because the interpolated grid lines are distorted by the rearrangement of landmarks relative to one another, warped transformation grids do not display any changes directly attributable to specific landmarks. That is true for the warped transformation grid per se (as in Fig. 2b), but not for some widely used ways to visualize transformation grids, with displacements of individual landmarks added to the transformation grid. Those landmark displacements, of course, do depend critically on the particular choice of icons and on the way the landmark configurations are superimposed on top of each other.

A crucial factor for the appearance and interpretation of transformation grids is the arrangement of the grid in the starting form, namely its alignment in relation to the anatomical axes (anterior–posterior, dorsal–ventral, and left–right) of the structure under study as well as the number and spacing of the grid lines. These choices affect the relations of the grid to the landmarks, for instance, whether specific landmark lies in the space between grid lines or whether a grid line passes through it. Changing the directions and spacing of the grid lines in the starting form, or the orientation of the grid in relation to the configuration, can therefore lead to substantial differences in the appearance of the warped

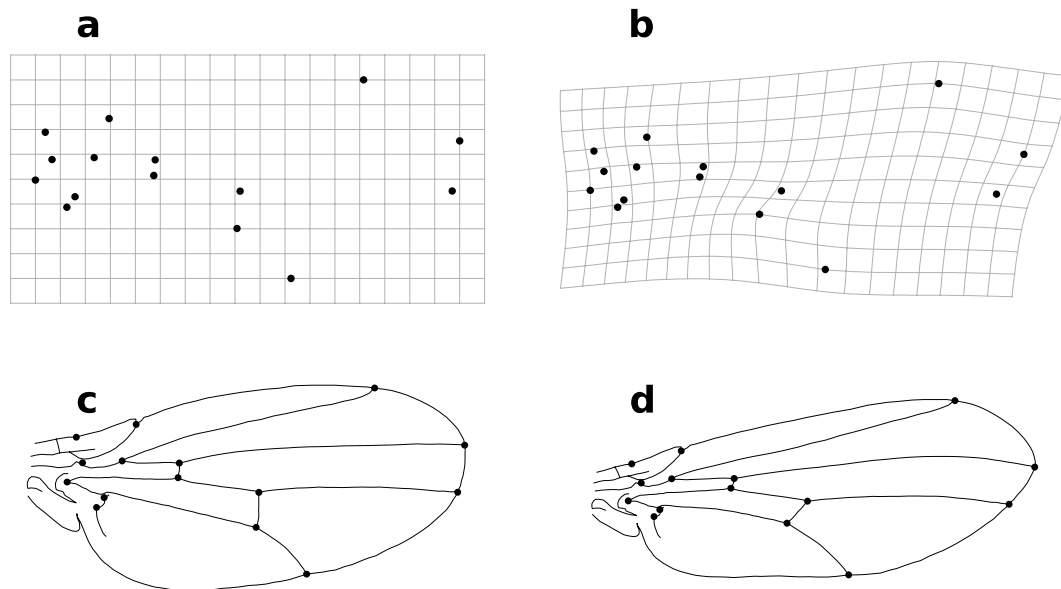


Fig. 2 Visualizations of a shape change in *Drosophila* wings with warped transformation grids and outline drawings. The shape change is the same from **a** to **b** as from **c** to **d**. By knowing that the starting grid is perfectly rectangular and that all grid cells are square, the

viewer can interpret a shape change based only on the warped grid **b** alone. For warped outline drawings, it is best to display graphs of the starting (**c**) and target shapes (**d**) side by side

transformation grid that visualizes the target shape, even for a single, constant shape change (Klingenberg 2013b). To some extent, these difficulties can be circumvented by using 2D drawings (Fig. 2c, d) or 3D surfaces instead of a rectangular grid, which provides a more intuitive anatomical interpretation (Klingenberg 2013b).

Transformation grids or warped outline drawings or surfaces, with starting and target shapes presented side by side (Fig. 2), can visualize shape changes without superimposing two icons. As a result, these graphs do not give the appearance that each landmark is shifting from one specific position to another. Therefore, these visualizations reduce the temptation for the viewers to ascribe shape changes to individual landmarks. Thereby, these visualizations also avoid some of the misconceptions related to the Pinocchio effect.

Simulation Studies and the Pinocchio Effect

Most published explanations of the Pinocchio effect use simulated data or examples in some way (Chapman 1990; Rohlf and Slice 1990; Walker 2000; Gill et al. 2007; Hallgrímsson et al. 2015; Tatsuta et al. 2018; Palci and Lee 2019), but they are widely used in different contexts as well, where they raise similar questions. Therefore, it is useful to think about simulations and the reasoning that underlies them. Simulation studies are special in science because the investigator can freely construct the underlying theoretical model that defines the true process generating the data. Also, the investigator has complete knowledge of the model. This is different from empirical scientific studies, where various models can be fit to the data, but where it is never completely known whether a particular model is actually true.

In the context of shape analysis, simulation studies are special in that the model also contains a specific set of icons, and therefore implies a particular size, position and orientation for each of the landmark configurations. This is clearest in simple simulations demonstrating the Pinocchio effect by comparing two landmark configurations that are identical except for a difference in the position of a single landmark (as in Fig. 1a; Chapman 1990; Walker 2000; Gill et al. 2007; Palci and Lee 2019) or of just a few landmarks (Siegel and Benson 1982; Hallgrímsson et al. 2015; Tatsuta et al. 2018). The fact that landmark configurations aligned by Procrustes superimposition differ in the positions of many landmarks, not just of those landmarks where the differences were in the simulated configurations, is then interpreted as a flaw of the Procrustes approach. The reasoning that underlies this interpretation has not been explained in detail or analyzed critically.

First, recall that Procrustes superimposition is a method for extracting shape variation. Accordingly, Procrustes-superimposed landmark configurations preserve the shapes

of the configurations in the original data, but the non-shape components of variation are standardized as part of the procedure. It is therefore clearly reasonable to expect Procrustes superimposition to recover shape differences between configurations. It is not nearly as clear, however, whether a method that is intended to extract only shape differences should reasonably be expected to recover a particular pair of icons including the possible differences in their sizes, positions or orientations. This raises the question whether the changes simulated in demonstrations of the Pinocchio effect are shape changes or whether they also involve changes in size, position and orientation.

For the example of Pinocchio's nose, it is fairly clear that moving the landmark at the tip of the nose away from the remaining landmarks on the face and head moves that landmark away from the overall centroid, and thereby in general will increase centroid size. This association of shape and size in the example of Pinocchio's nose has been acknowledged (Hallgrímsson et al. 2015). Similarly, for the example of Fig. 1a, it is clear that the black figure shares borders with the gray figure or is contained inside it, and therefore must have a smaller centroid size (and is also smaller in terms of its area and other size measures involving all five landmarks). These examples show that the non-shape aspects of a landmark configuration can be affected, but they raise the more general question whether shifts of only one or a few landmarks always involve the non-shape components of size, position, and orientation, or whether there are such changes that affect only shape. To answer this question, it is necessary to consider whether it is possible to change the position of one or a few landmarks in a configuration while maintaining the various constraints on shape variation (Klingenberg 2020).

For shifts in a single landmark, it is possible to draw some clear and general conclusions by considering what type of small shifts in single landmarks can maintain specific aspects of the configuration. To maintain the same centroid size for the whole configuration, a landmark can only move on a circle around the centroid of the configuration, but it cannot shift in a radial direction toward or away from the centroid (Fig. 3a). To maintain the same overall orientation for the configuration, a landmark can move on a straight line toward or away from the centroid (Fig. 3b). Combining these two constraints means that there is no way for a single landmark to change its position that simultaneously maintains a constant centroid size and constant orientation for the overall configuration. Any shift in a single landmark also inevitably alters the position of the centroid of the configuration, which implies a shift in the overall position of the configuration. To compensate for this, a change of every other landmark in the opposite direction would be required (Fig. 3c). Overall, therefore, it is clear that there cannot be a change in the position of a single landmark in a configuration that yields

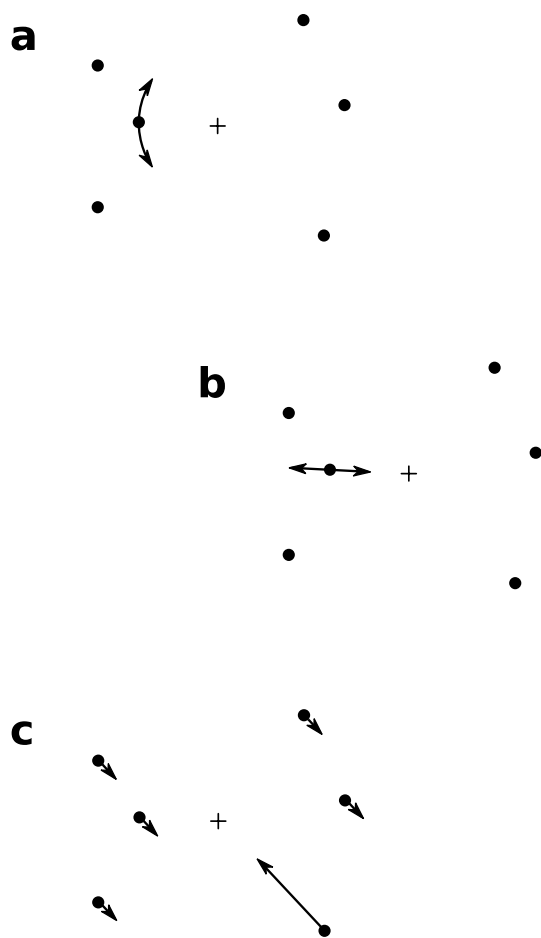


Fig. 3 Changes to single landmarks in a configuration. Each graph shows a configuration of six landmarks (black dots) and their centroid (+). All landmark shifts are exaggerated for better visibility. **a** To preserve a constant centroid size of the configuration, a landmark can move on a circle around the centroid. **b** To maintain the orientation of the configuration, a landmark can move along a straight line from the centroid. **c** To maintain a constant position for the configuration, any shift of a landmark requires a smaller shift of all the remaining landmarks in the opposite direction

only a change of shape, but leaves size, position and orientation unaltered. Simulations of the Pinocchio effect based on shifts of single landmarks therefore inevitably entail changes in the size, position or orientation of the landmark configuration.

It is more difficult to gauge the consequences of shifts in two or more landmarks because the effects of changes in one landmark potentially can be balanced by changes in other landmarks. Nevertheless, it is difficult to think of combinations of changes in just two landmarks so that all the shape constraints are met for the whole landmark configuration. Also, based on the nature of those constraints (Klingenberg 2020), it is likely that most or all the landmarks, rather than just a small number, are involved in most shape changes. The requirement can also be expressed algebraically: the

vector of the shape change must be orthogonal to all the four (for 2D data) or seven (for 3D data) vectors characterizing the shape constraints, which in part depend on the starting shape (Klingenberg 2020). This condition is challenging to meet. Changes that are concentrated in just a few landmarks are very likely to include components of change in the non-shape features of size, position, and orientation. Therefore, it generally cannot be expected that methods that focus of shape variation, such as Procrustes superimposition, can recover the changes in simulated examples of changes in one or only a few landmarks.

A different type of simulation generates samples of landmark configurations by using a specific configuration as the mean and adding variation as random displacements of landmarks drawn from some specified distribution. This approach has been used widely in the development of statistical shape analysis and is often called offset or perturbation model (Goodall 1991; Rohlf 2000, 2003; Lele and Richtsmeier 2001; Dryden and Mardia 2016). Some of these simulation studies have used isotropic normal displacements (Rohlf 2000, 2003; Cardini 2019), where deviations for all landmark coordinates are drawn independently from normal distributions with zero means and the same variance, distributions with various degrees of covariation between landmarks (Goswami et al. 2019), or a distribution so that some landmarks are more variable than others (Walker 2000). The simulated data then undergo a Procrustes superimposition and usually some further analyses. Often the results of these analyses are compared to the original simulated data, or at least the results are interpreted in relation to how the data were simulated (Walker 2000; Cardini 2019; Goswami et al. 2019). For instance, Walker (2000) found that the covariance matrix of Procrustes-superimposed landmark configurations differed from the covariance matrix of the original simulated configurations, more so than the covariance matrix derived from configurations superimposed by generalized resistant fit (Rohlf and Slice 1990). Cardini (2019) observed that, after Procrustes superimposition, there was covariation among landmarks and that this covariation was structured to suggest modularity, even though the data were simulated from isotropic distributions and independently among landmarks. Similarly, Goswami et al. (2019) ran simulations using covariance structures with various patterns of modularity and found that most but not all of them could be recovered from Procrustes-superimposed landmark coordinates. All these studies discussed the results as changes or distortions of the covariance structure induced by Procrustes superimposition. None of these studies, however, paid explicit attention to the fact that the difference between covariance structures of simulated data before and after Procrustes superimposition also involves the difference that the original simulated data include variation in both shape and the non-shape components of size, position and orientation,

whereas Procrustes-superimposed data contain exclusively shape variation.

Some of the observed effects in these studies result from the difference between the unconstrained variation of landmarks in the simulated data and the Procrustes-superimposed data that must obey the shape constraints (Klingenberg 2020). One of the main results of Walker's (2000) study is that the discrepancies between methods were largest for small numbers of landmarks, and decreased as more landmarks were included. To understand this effect, it is helpful to consider the dimensionality of the data. For 2D data with k landmarks, there are $2k$ landmark coordinates and four shape constraints, resulting in $2k - 4$ shape dimensions that are retained in the Procrustes-superimposed data. With an isotropic distribution of landmark displacements, every dimension of the original variation contains the same amount of variation: if there are $2k$ units of variation overall, of which $2k - 4$ units are shape variation. It follows that the proportion of the non-shape variation that is removed by the Procrustes fit is $4/2k$. This proportion is quite large if few landmarks are considered (e.g., it is $2/3$ for triangles), but it diminishes with increasing numbers of landmarks. Because Walker's (2000) simulations did not use isotropic variation, these calculations don't apply directly to his study, but the same effects still are bound to produce some reduction from the unconstrained variation in the simulated landmark deviations to exclusively shape variation in the Procrustes-superimposed configurations, and the calculation for the isotropic data explains the finding that the relative discrepancies diminish with increasing numbers of landmarks.

The changes in the covariance patterns before and after Procrustes superimposition that were observed in some simulation studies (Cardini 2019; Goswami et al. 2019) can also be understood as consequences of the shape constraints (Klingenberg 2020). Because these constraints almost always involve all landmarks, applying them to simulated data may change covariation among landmarks. Translations always affect all landmarks equally, and corresponding constraints may therefore also have joint effects on landmarks. The vectors for scaling and rotation are functions of the reference shape (Klingenberg 2020), which is best chosen in the present as the landmark configuration used to simulate data. Accordingly, these constraints may affect the landmarks differentially, but they are also likely to affect all or most landmarks somehow. It is therefore not surprising that the covariance structure changes when the shape constraints are applied.

These considerations have major implications for how the results of simulation studies should be interpreted. In particular, they call into question the implicit assumption that the covariance structure of Procrustes-superimposed landmark configurations should match the covariance structure of the same configurations as they were generated in the

simulations. The data simulated by random displacements of landmarks around a given starting configuration contain variation in shape as well as variation in the non-shape features of size, position and orientation. By contrast, Procrustes-superimposed landmark configurations contain only shape variation. A discrepancy between the covariance structures of the original and Procrustes-superimposed landmark configurations is not an indication of a poor performance of the Procrustes method, but simply a consequence of whether or not the non-shape components of variation are included or not. If the goal is to find a method that represents the covariance structure of the original data, Procrustes superimposition is inherently a poor choice of method because it focuses on shape only and omits all non-shape components of variation. If the goal of the simulation is to extract shape variation, discrepancies between the simulated data and the Procrustes-superimposed configurations are not a problem. Instead, the investigator might want to consider using a different approach for the simulation, which provides pure shape variation without an additional non-shape component.

Natural Space and Natural Superimposition: How Natural are They?

The ideas implied in the visualizations of superimposed shapes and in simulations of landmark variation using offset models have been formalized in the closely related concepts of the "natural space" (Richtsmeier et al. 2005) and the "natural superimposition" (Goswami et al. 2019). Both of these embody a notion that the non-shape parameters of the models, namely the position, orientation and possibly size of landmark configurations included in comparisons or produced in simulations, are in some way natural in relation to the scientific context of morphometric studies. The motivation clearly comes from the offset shape models, within which the underlying space is indeed natural in relation to the generating process, but there is a perception that these concepts also apply beyond these models.

The concept of the natural space was proposed with little explanation except that this is the coordinate system where within-sample variation arises (Richtsmeier et al. 2005), whereas Goswami et al. (2019, p. 679) stated explicitly that the concept of the natural superimposition "is a biologically vague idea". To understand some of the reasoning underlying these concepts in more detail, it is helpful to consider the fictitious examples offered by Richtsmeier and colleagues (Lele and Richtsmeier 2001; Richtsmeier et al. 2002) to explain similar ideas before they coined the concept of the natural space (Richtsmeier et al. 2005). One example involves a landmark configuration that is drawn with red marker pen on a transparency (Lele and Richtsmeier 2001). Additional transparencies are then laid on top of it (without

any effort to align the margins of the transparencies) and the original landmarks are copied onto these, with a small amount of variability, with a black pen. After this, someone drops the pile of transparencies on the floor, and the transparency with the red marks is lost altogether. As a result, the landmark configurations are no longer in the original alignment to each other. Reconstructing precisely the same alignment and starting configuration would require knowing the exact displacements of the transparencies. Alternatively, one might try to superimpose the transparencies in a way so that the alignment of the landmark configurations is optimal in some sense, and to find a consensus to reconstruct the starting configuration. The second example is a machine that produces wax figurines of galagos using a mold (Richtsmeier et al. 2002). After the wax galagos are taken from the mold, various anatomical details are subtly modified by hand and the figures are placed on a shelf in haphazard location and orientation. For some reason, the mold in the machine is lost, and the operator wants to reconstruct a mold from the wax figures that were previously produced. This would require superimposing multiple wax galagos to produce some consensus from which the mold could be reconstructed. The original alignment of transparencies and the mold stand for the idea of a true position and orientation that the later publication (Richtsmeier et al. 2005) called the natural space.

To evaluate how robust the conclusions from these examples are for application to real biological data, we can try to modify the examples slightly and see whether they still give rise to the idea of a natural space. Instead of tracing the landmark configuration from the first transparency directly to the others, let's imagine that we project the original transparency onto the screen in a classroom, hand out blank transparencies and black pens to the students, and ask them to copy the projected landmark configuration onto their transparencies. We can then collect all transparencies together and overlay them to find a consensus and to characterize the variation among the students' versions. For the second example, let's assume that, instead of producing wax galagos in a machine using a mold, there is an artist who models or carves each figurine separately as copies of an original galago sculpture. These modifications of the two scenarios differ in some respects from the original examples. First, variation in size is an obvious additional factor that comes into play and can enter the analyses. Second, the processes that produce the landmark configurations or galago figurines involve no alignment that is original or special in any sense. Each new copy of the landmark configuration or each new galago figure has its own location and coordinate system, and there is no position or coordinate system that is privileged by comparison to others.

For the published versions of the hypothetical examples (Lele and Richtsmeier 2001; Richtsmeier et al. 2002), the

location of the original transparency with the landmarks drawn in red pen or of the mold in the galago example is special because the new copies originate in that specific location and orientation. For this reason, this special coordinate system is closely related to the idea of a natural space (Richtsmeier et al. 2005). By contrast, for my modified versions above, there is no such privileged location or coordinate system. In that version, there is no equivalent whatsoever for a space or superimposition that would be any more natural than any other. From this comparison between the two versions of the examples, it becomes clear that the fundamental insight the examples were intended to convey depends on the details of the process used in each example: it requires tracing the landmark configurations directly from one transparency to another one laid on top of it rather than copying by eye from a screen to a transparency on the table, or producing galago figures by molding rather than modeling. If a small change in the imaginary scenario makes such a big difference for the conclusion that follows from it, this casts doubt on whether the notion of some coordinate systems being special or more natural than others is justified at all.

This reasoning is invalid, of course, if my modifications of the scenarios are misrepresentations of the biological processes responsible for the variation in morphometric data. This raises the question whether the original scenarios or my modified versions are a more appropriate (not to mention realistic) analogue of the processes by which biological structures originate. In other words, is there some direct transfer of morphological form, such as it is characteristic for direct tracing of landmark configurations or molding of wax figurines, or are the processes of copying or modeling in a different place perhaps better analogues for biological processes? Developmental biology indicates that the latter is the case. Morphological structures arise anew without direct contact with a preexisting copy of the corresponding structure. For instance, mammalian skulls develop in the embryo in the maternal uterus, with no direct contact to the maternal skull. Even more generic features such as the anatomical axes of an embryo (anterior–posterior, dorsal–ventral and left–right) in many animals originate in the egg or embryo (e.g., by gravity, via the site of sperm entry into the egg, etc.) and are often not carried over from the maternal body axes (Goldstein and Freeman 1997). Direct transfer of features is rare in the development of organisms, and even where it occurs, it is often limited to cellular components such as cell membrane, which undergoes fundamental changes before the adult condition is reached. Overall, therefore, my modified versions of the fictitious examples of the landmarks on transparencies (Lele and Richtsmeier 2001) or of the galago figurines (Richtsmeier et al. 2002) may actually be rather closer to biological reality than the original versions. As a result, those examples do nothing to justify the idea of a true or natural alignment among landmark configurations when it comes to morphometric analyses of biological data:

there is nothing natural about the “natural space” or “natural superimposition”.

There is a context, however, where the natural space and natural superimposition are perfectly natural: simulations using the perturbation model (Rohlf and Slice 1990; Goodall 1991; Lele and Richtsmeier 2001; Richtsmeier et al. 2005; Goswami et al. 2019). Within this model, the original alignment in which a sample of landmark configurations is generated from a starting configuration is clearly natural. Each landmark configuration generated in the sample is a particular icon that has not only a particular shape, but also a particular size, position, and orientation. This icon is natural because of how it originated in the perturbation model. For empirical landmark data, by contrast, no such original alignment of landmark configurations exists. All icons of the equivalence class that corresponds to a particular shape are equivalent, and none is any more or any less natural than any other. Richtsmeier et al. (2002) concluded that “in practice it is unrealistic to try to determine the true orientation of a biological form” (p. 69). This conclusion is justified, but it is not just a practical issue for empirical studies, but it goes deeper, because no single true orientation exists for biological objects. The positions and orientations of the specimens from which landmark data are collected are irrelevant for the biological questions that morphometric studies address. Investigators are free to place specimens in any relation to the camera, scanner or digitizer that is convenient for the use of that equipment. Extraction of shape or form information from the data automatically removes the arbitrary variation in position and orientation.

What does this mean for the Pinocchio effect? These arguments further emphasize the importance to keep in mind that shape variation is inherently relative in terms of how variation can be allocated to specific landmarks by the choice of particular icons. The Pinocchio effect is usually invoked if a shape change can be characterized by a pair of icons that differ only (or at least mostly) in the position of a single landmark, whereas the remaining landmarks coincide (e.g., Fig. 1a). As shown above, this alignment of landmark configurations is no more natural or true than any other alignment involving the same equivalence classes of icons (e.g., Figs. 1b, c). And just as with the hypothetical examples of the landmark configurations drawn on transparencies or of the galago figurines, simulations of the Pinocchio effect using variants of the perturbation model (Rohlf and Slice 1990; Walker 2000) do not realistically match the situation of empirical studies of shape.

Summary and Outlook

The Pinocchio effect has a long and undistinguished history of confusing morphometricians, including myself, who have referred to it for explaining results of empirical analyses,

computer simulations or theoretical considerations (Chapman 1990; Klingenberg and McIntyre 1998; Walker 2000; Slice 2005; Gill et al. 2007; von Cramon-Taubadel et al. 2007; Zelditch et al. 2012; Hallgrímsson et al. 2015; Fruciano 2016; Tatsuta et al. 2018; Palci and Lee 2019). Given that most morphometric studies aim to investigate variation in shape, which is defined as all the geometric features of an object except for its size, position and orientation, it is ironic that the concerns about the Pinocchio effect focus precisely on those aspects, position and orientation, that are not aspects of shape.

The Pinocchio effect stems from the widespread preference for superimpositions that involve change in just one or few landmarks in a configuration over superimpositions where many landmarks change positions (e.g., Fig. 1a–c). Visualizations that use pairs of icons where changes are concentrated in few landmarks tend to appear more “natural” or “correct” than those where changes are distributed over many landmarks, even if they actually involve the same shape difference. This is a powerful perceptual bias that has considerably influenced the thinking about superimposition methods and especially about the Pinocchio effect. The reasoning presented in this paper makes it clear that there is no basis for any such preference in the logic that underlies the methods of geometric morphometrics.

Perhaps the greatest inherent difficulty of morphometric methods is the fact that it is not possible to assign shape changes unambiguously to individual landmarks or even particular sets of landmarks. To the contrary, the relative nature shape changes encompasses all the landmarks in the configuration under study. There are far-reaching consequences of this realization that extend beyond the Pinocchio effect, which are especially relevant for studies of morphological integration and modularity (Klingenberg 2013a; Cardini 2019; Goswami et al. 2019). Those consequences need to be fully explored elsewhere.

Compliance with Ethical Standards

Conflict of interest The author has no conflict of interest.

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

References

- Albert, M. H., Le, H., & Small, C. G. (2003). Assessing landmark influence on shape variation. *Biometrika*, *90*, 669–678.
- Avery, G. S., Jr. (1933). Structure and development of the tobacco leaf. *American Journal of Botany*, *20*(9), 565–592.
- Boas, F. (1905). The horizontal plane of the skull and the general problem of the comparison of variable forms. *Science*, *21*, 862–863.
- Bookstein, F. L. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 567–585.
- Bookstein, F. L. (1996). Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology*, *58*, 313–365.
- Cardini, A. (2019). Integration and modularity in Procrustes shape data: Is there a risk of spurious results? *Evolutionary Biology*, *46*, 90–105.
- Chapman, R. E. (1990). Conventional Procrustes approaches. In F. J. Rohlf, & F. L. Bookstein (Eds.), *Proceedings of the Michigan morphometrics workshop* (pp. 251–267). Ann Arbor, MI: University of Michigan Museum of Zoology.
- Courtenay, L. A., Herranz-Rodrigo, D., Huguet, R., Maté-González, M. A., González-Aguilera, D., & Yravedra, J. (2020). Obtaining new resolutions in carnivore tooth pit morphological analyses: A methodological update for digital taphonomy. *PLoS ONE*, *15*(10), e0240328.
- Dryden, I. L., & Mardia, K. V. (2016). *Statistical shape analysis, with applications in R* (2nd ed.). Chichester: Wiley.
- Fox, N. S., Veneracion, J. J., & Blois, J. L. (2020). Are geometric morphometric analyses replicable? Evaluating landmark measurement error and its impact on extant and fossil *Microtus* classification. *Ecology and Evolution*, *10*, 3260–3275.
- Fruciano, C. (2016). Measurement error in geometric morphometrics. *Development Genes and Evolution*, *226*, 139–158.
- Gill, D., Ritov, Ya., & Dror, G., et al. (2007). Is Pinocchio's nose long or his head small? Learning shape distances for classification. In G. Bebis, R. Boyle, B. Parvin, D. Koracin, N. Paragios, & S.-M. Tancveer (Eds.), *Advances in Visual Computing* (pp. 652–661). Berlin: Springer.
- Gilmour, D., Rembold, M., & Leptin, M. (2017). From morphogen to morphogenesis and back. *Nature*, *541*, 311–320.
- Goldstein, B., & Freeman, G. (1997). Axis specification in animal development. *BioEssays*, *19*(2), 105–116.
- Goodall, C. R. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society B*, *53*, 285–339.
- Goswami, A., Watanabe, A., Felice, R. N., Bardua, C., Fabre, A.-C., & Polly, P. D. (2019). High-density morphometric analysis of shape and integration: The good, the bad, and the not-really-a-problem. *Integrative and Comparative Biology*, *59*(3), 669–683.
- Granier, C., & Tardieu, F. (1998). Spatial and temporal analyses of expansion and cell cycle in sunflower leaves. *Plant Physiology (Rockville)*, *116*, 991–1001.
- Hallgrímsson, B., Percival, C. J., Green, R., Young, N. M., Mio, W., & Marcucio, R. (2015). Morphometrics, 3D imaging, and craniofacial development. *Current Topics in Developmental Biology*, *115*, 561–597.
- Heisenberg, C.-P., & Bellaïche, Y. (2013). Forces in tissue morphogenesis and patterning. *Cell*, *153*, 948–962.
- Kendall, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, *16*, 81–121.
- Kendall, D. G., Barden, D., Carne, T. K., & Le, H. (1999). *Shape and shape theory*. Chichester: Wiley.
- Kennaway, R., Coen, E., Green, A., & Bangham, A. (2011). Generation of diverse biological forms through combinatorial interactions between tissue polarity and growth. *PLoS Computational Biology*, *7*, e1002071.
- Kitching, I. J., Forey, P. L., Humphries, C. J., & Williams, D. M. (1998). *Cladistics: The theory and practice of parsimony analysis* (2nd ed.). Oxford: Oxford University Press.
- Klingenberg, C. P. (2013a). Cranial integration and modularity: Insights into evolution and development from morphometric data. *Hystrix-Italian Journal of Mammalogy*, *24*(1), 43–58.
- Klingenberg, C. P. (2013b). Visualizations in geometric morphometrics: How to read and how to make graphs showing shape changes. *Hystrix-Italian Journal of Mammalogy*, *24*(1), 15–24.
- Klingenberg, C. P. (2016). Size, shape, and form: Concepts of allometry in geometric morphometrics. *Development Genes and Evolution*, *226*(3), 113–137.
- Klingenberg, C. P. (2020). Walking on Kendall's shape space: Understanding shape spaces and their coordinate systems. *Evolutionary Biology*. <https://doi.org/10.1007/s11692-11020-09513-x>
- Klingenberg, C. P., & McIntyre, G. S. (1998). Geometric morphometrics of developmental instability: Analyzing patterns of fluctuating asymmetry with Procrustes methods. *Evolution*, *52*, 1363–1375.
- Lele, S. R., & Richtsmeier, J. T. (2001). *An invariant approach to statistical analysis of shapes*. Boca Raton: Chapman & Hall/CRC.
- Montell, D. J. (2008). Morphogenetic cell movements: Diversity from modular mechanical properties. *Science*, *322*, 1502–1505.
- Palci, A., & Lee, M. S. Y. (2019). Geometric morphometrics, homology and cladistics: Review and recommendations. *Cladistics*, *35*, 230–242.
- Poethig, R. S., & Sussex, I. M. (1985). The developmental morphology and growth dynamics of the tobacco leaf. *Planta*, *165*, 158–169.
- Richtsmeier, J. T., DeLeon, V. B., & Lele, S. R. (2002). The promise of geometric morphometrics. *Yearbook of Physical Anthropology*, *45*, 63–91.
- Richtsmeier, J. T., Lele, S. R., & Cole, T. M. (2005). Landmark morphometrics and the analysis of variation. In B. Hallgrímsson & B. K. Hall (Eds.), *Variation: a central concept in biology* (pp. 49–69). Burlington: Elsevier.
- Rohlf, F. J. (2000). Statistical power comparisons among alternative morphometric methods. *American Journal of Physical Anthropology*, *111*, 463–478.
- Rohlf, F. J. (2003). Bias and error in estimates of mean shape in geometric morphometrics. *Journal of Human Evolution*, *44*, 665–683.
- Rohlf, F. J., & Slice, D. E. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, *39*, 40–59.
- Sagner, A., & Briscoe, J. (2017). Morphogen interpretation: Concentration, time, competence, and signaling dynamics. *WIREs Developmental Biology*, *6*, e271.
- Saurer, W., & Possingham, J. V. (1970). Studies on the growth of spinach leaves (*Spinacea oleracea*). *Journal of Experimental Botany*, *21*(66), 151–158.
- Schmidt-Nielsen, K. (1984). *Scaling: Why is animal size so important?* Cambridge: Cambridge University Press.
- Siegel, A. F., & Benson, R. H. (1982). A robust comparison of biological shapes. *Biometrics*, *38*, 341–350.
- Slice, D. E. (2005). Modern morphometrics. In D. E. Slice (Ed.), *Modern Morphometrics in Physical Anthropology* (pp. 1–45). New York: Kluwer Academic/Plenum.
- Sneath, P. H. A. (1967). Trend-surface analysis of transformation grids. *Journal of Zoology*, *151*, 65–122.
- Tatsuta, H., Takahashi, K. H., & Sakamaki, Y. (2018). Geometric morphometrics in entomology: Basics and applications. *Entomological Science*, *21*, 164–184.
- Thompson, D. A. W. (1961). *On growth and form*. Cambridge: Cambridge University Press.

- Thulman, D. K. (2019). A typology of Florida fluted points using landmark-based geometric morphometrics. *Paleoamerica*, 2019(2), 181–190.
- Varón-González, C., Whelan, S., & Klingenberg, C. P. (2020). Estimating phylogenies from shape and similar multidimensional data: Why it is not reliable. *Systematic Biology*, 69(5), 863–883.
- von Cramon-Taubadel, N., Frazier, B. C., & Lahr, M. M. (2007). The problem of assessing landmark error in geometric morphometrics: Theory, methods, and modifications. *American Journal of Physical Anthropology*, 134, 24–35.
- Walker, J. A. (2000). Ability of geometric morphometric methods to estimate a known covariance matrix. *Systematic Biology*, 49, 686–696.
- Wang, L., Beyer, S. T., Cronk, Q. C. B., & Walus, K. (2011). Delivering high-resolution landmarks using inkjet micropatterning for spatial monitoring of leaf expansion. *Plant Methods*, 7, 1.
- Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric morphometrics for biologists: A primer* (2nd ed.). Amsterdam: Elsevier.

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