The Problem of Assessing Landmark Error in Geometric Morphometrics: Theory, Methods and Modifications

N. von Cramon-Taubadel[1], B. C. Frazier[2] and M. Mirazón Lahr[1]

[1]Leverhulme Centre for Human Evolutionary Studies, University of Cambridge.

[2]Department of Anthropology, Penn State University

Introduction. Geometric morphometric methods rely on the accurate identification and quantification of landmarks on biological specimens. A review of methods currently being employed to assess measurement error identified three general approaches to the problem. Shortcomings of these methods are reviewed and discussed here. A revised method is proposed and described that mitigates the problems associated with existing methods.

Error assessment at the configuration level

• Some studies assess the effect of measurement error on a landmark configuration by examining repeat measurements of specimens in shape space. Following superimposition, repeats of a single specimen should cluster together and be distinguishable from other specimens when tested in a principal components analysis[1,2] (Fig.1a).

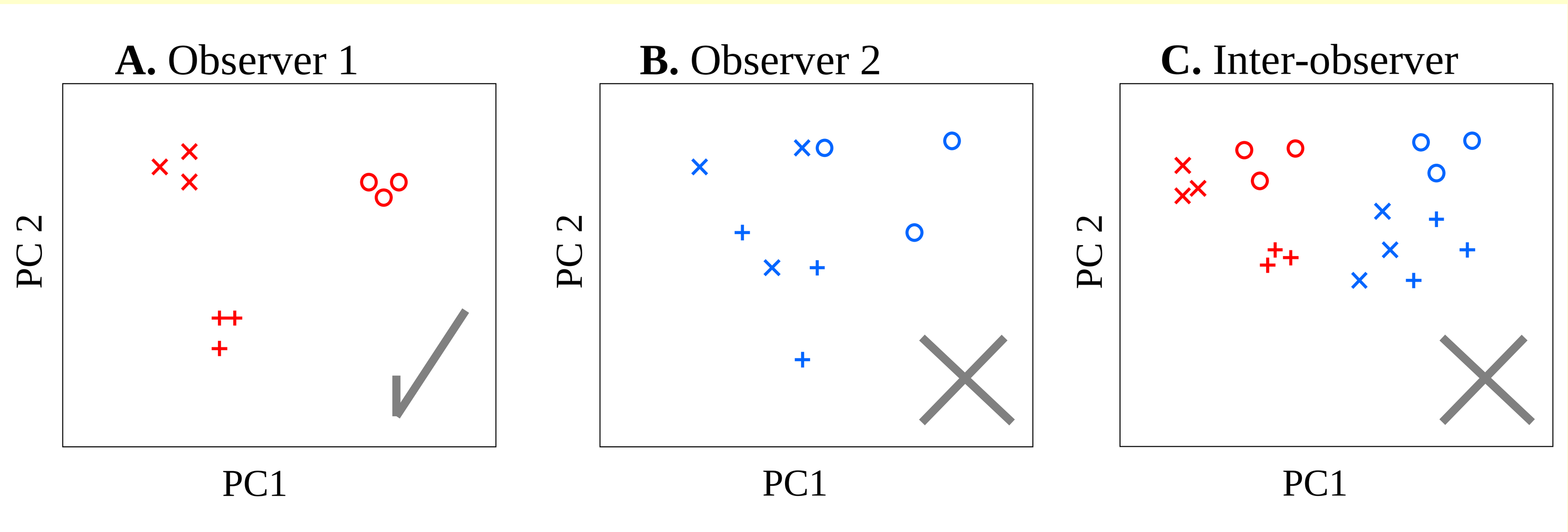

Figure 1. Three measurement repeats from three different specimens are compared in a principal components analysis. Each mark represents the location of a single repeat of a full landmark configuration in shape space. **A**. Repeats of each specimen clearly cluster. **B**. Repeats of a given specimen do not cluster. **C**. Repeats of the same specimen cluster separately for two observers.

• If repeats do not cluster tightly (Fig. 1b), relatively error-prone landmarks cannot be identified. Additionally, two observers may take landmarks at slightly different locations, which could cause the primary source of shape variation to be inter-observer measurement variance rather than inter-specimen variance (Fig. 1c).

• Measurement error may only reside with a few landmarks, but this increased variance is transferred to the whole configuration during superimposition of repeats using Generalized Procrustes Analysis (GPA) – a phenomenon known as the "Pinocchio effect"[3] (Fig. 2). For this reason, if a researcher is concerned with the relative validity of individual landmarks, measurement error must be assessed on a landmark-by-landmark basis.

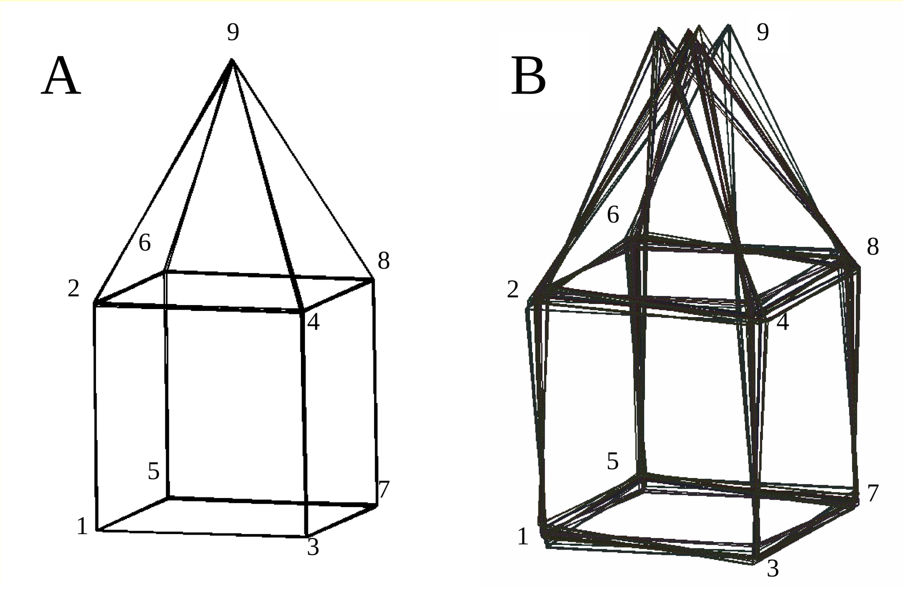

Figure 2. Pinocchio effect
A landmark configuration was repeatedly digitized such that only landmark 9 (the tip) varied in position relative to the other eight landmarks (**A**). Following superimposition of the configurations, the variation around landmark 9 was randomly transferred across the entire configuration (**B**).

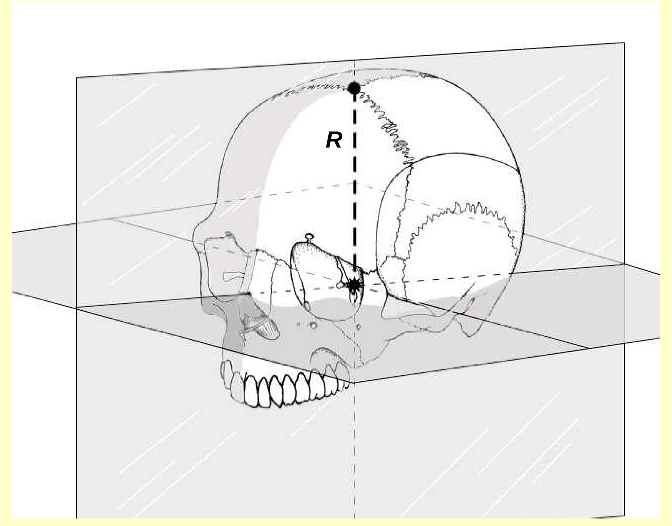

Figure 3. Example of the centroid radius (*R*) from Bregma.

Error assessment at the landmark level

• One suggestion[4] for assessing measurement error on an individual landmark basis involves the use of the *centroid radius,* the distance between any individual landmark and the configuration centroid (Fig. 3).

• The centroid radius corresponding to each landmark is calculated for each repeat, and the mean and percentage error associated with each centroid radius is computed across repeats.

• This method may potentially underestimate error, depending on the geometry of the landmark configuration (Fig. 4).

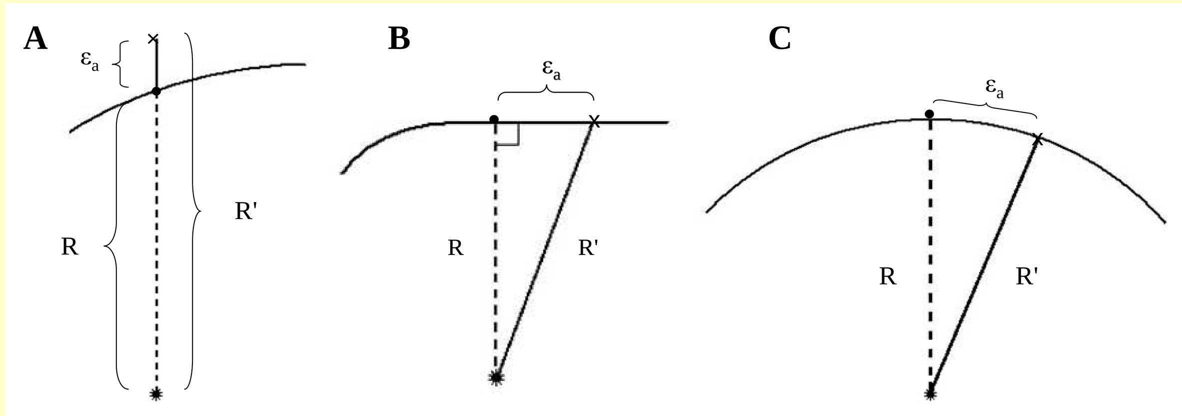

Figure 4. Potential discrepancies between the calculated error ($R' - R$) and the actual error (ε_a) when using the centroid radius (R) as a means of error assessment. **A.** If error occurs normal to the surface, R' and R are collinear and therefore $\varepsilon_a = R' - R$. **B.** Modeling the vicinity of the landmark as a flat surface, error (ε_a) occurs perpendicular to R. Thus, $R' = \sqrt{R^2 + \varepsilon_a^2}$. **C.** Modeled as a sphere, $R' = R$ and therefore no error along the surface of the specimen is detectable.

Error assessment while keeping a constant orientation

• Corner et al. (1992)[5] describe a method for assessing individual landmark precision by repeatedly digitizing landmarks while keeping the digitizer and the specimen in a constant orientation relative to each other (Fig. 5). The raw repeat values of each landmark can be directly compared for variation.

• Although this approach accurately measures landmark error, it can only be applied to landmarks taken on an object held static relative to a digitizer, or one that is in a constant frame of reference (such as a CT scan).

• However, many morphometric studies include instrumentally defined landmarks. In these cases, the error associated with the procedure used to identify the point on a specimen (Fig. 6) is intrinsic to the definition of the landmark in space.

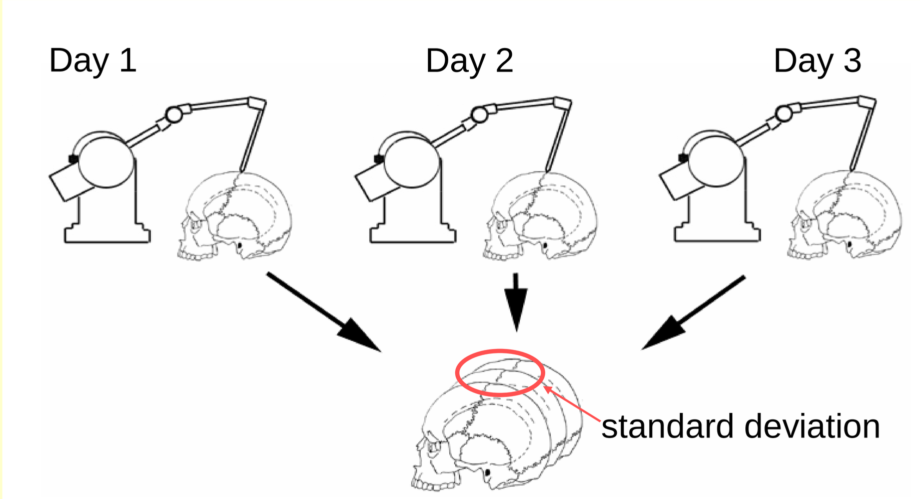

Figure 5. Corner et al. (1992) method for assessing landmark precision.

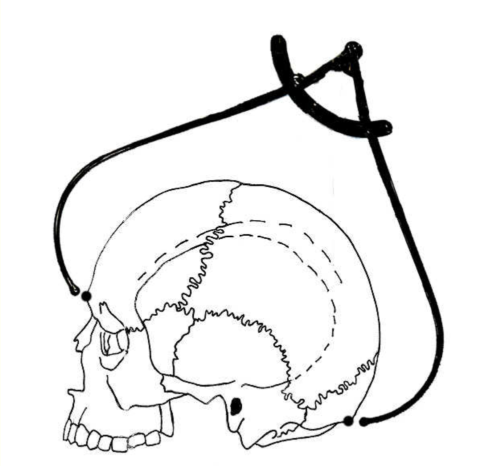

Figure 6. Examples of instrumentally defined landmarks: Glabella and Opistocranium.

Error assessment using partial superimposition (PS)

• We propose a method for assessing individual landmark precision using three *reference landmarks* to superimpose repeat trials, thus allowing specimens to be moved between measuring sessions.

• These reference landmarks must first be tested for precision using the Corner et al. method and must be determined to have low and approximately equal variances and to be uncorrelated.

• The precision of the remaining landmarks can then be evaluated on a landmark-by-landmark basis (Fig. 7).

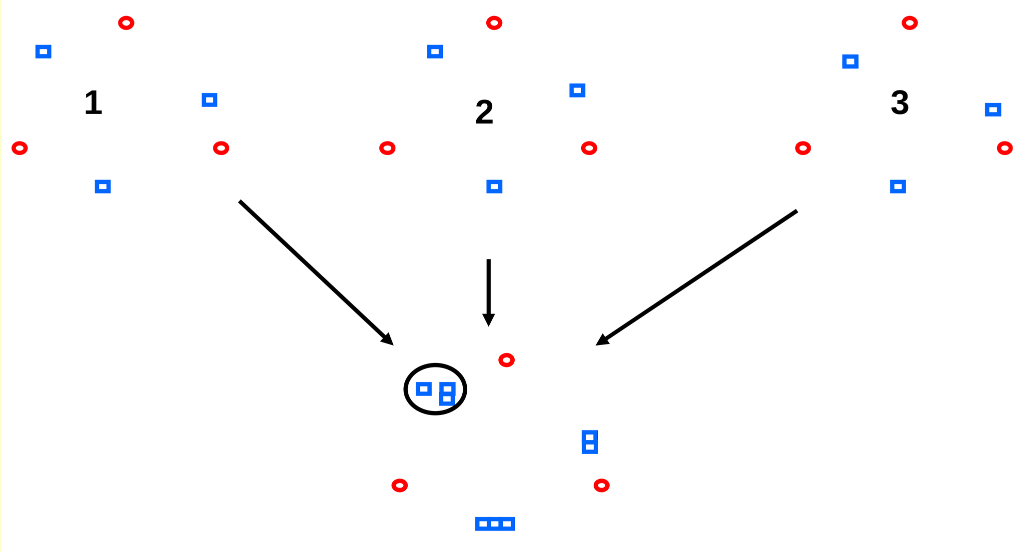

Figure 7. Partial superimposition of three repeat trials of a landmark configuration. Reference landmarks (red circles) are used to superimpose trials, such that measurement precision can be estimated for other landmarks (blue squares).

• A configuration of fourteen well-defined landmarks was assessed for intra-observer error using three methods: the Corner et al. (1992) method, PS as proposed here, and full superimposition using GPA.

• A plot of the calculated errors using the three methods (Fig. 8) demonstrates that PS produces the same *pattern* of results as Corner et al., in contrast to using GPA, which yields a random pattern of error estimates.

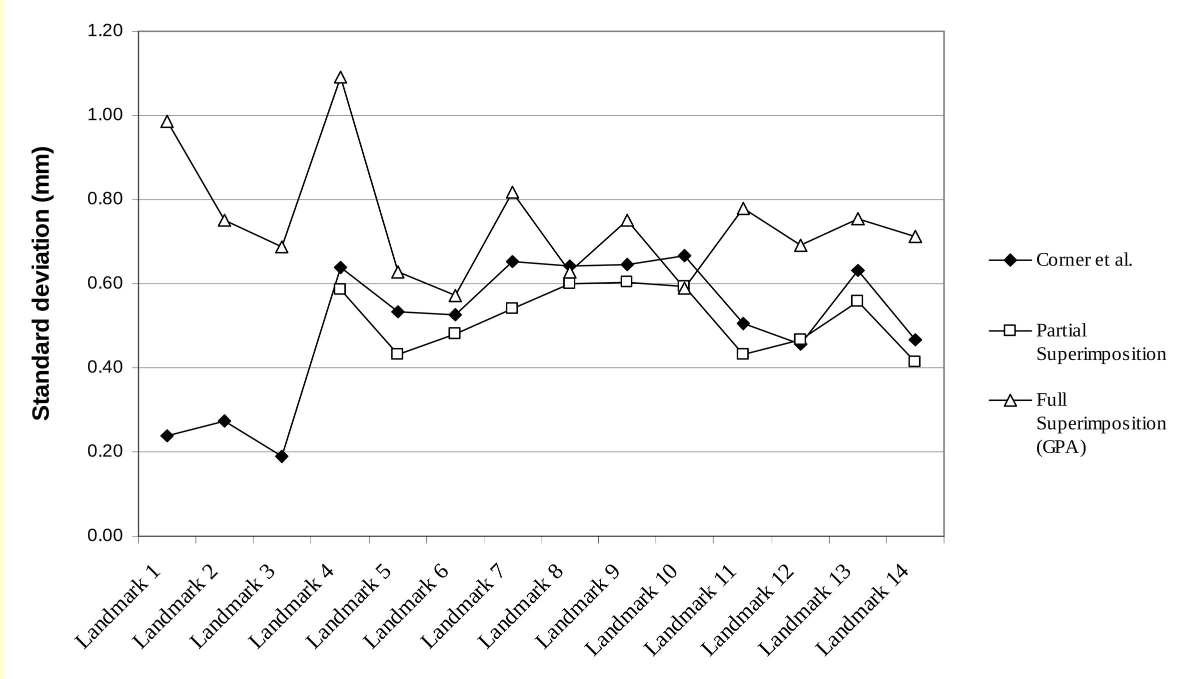

Figure 8. Comparison of three approaches to error assessment.

Making the transition from traditional to geometric morphometrics

• One potential difficulty in making the transition from traditional morphometrics to landmark-based morphometrics in anthropology is the inadequacy of the anatomical landmark definitions currently available.

• Describing a landmark in terms of a junction of sutures may provide an accurate *description* of an anatomical locality, but may fail to *define* a single point with sufficient exactness (Fig. 9).

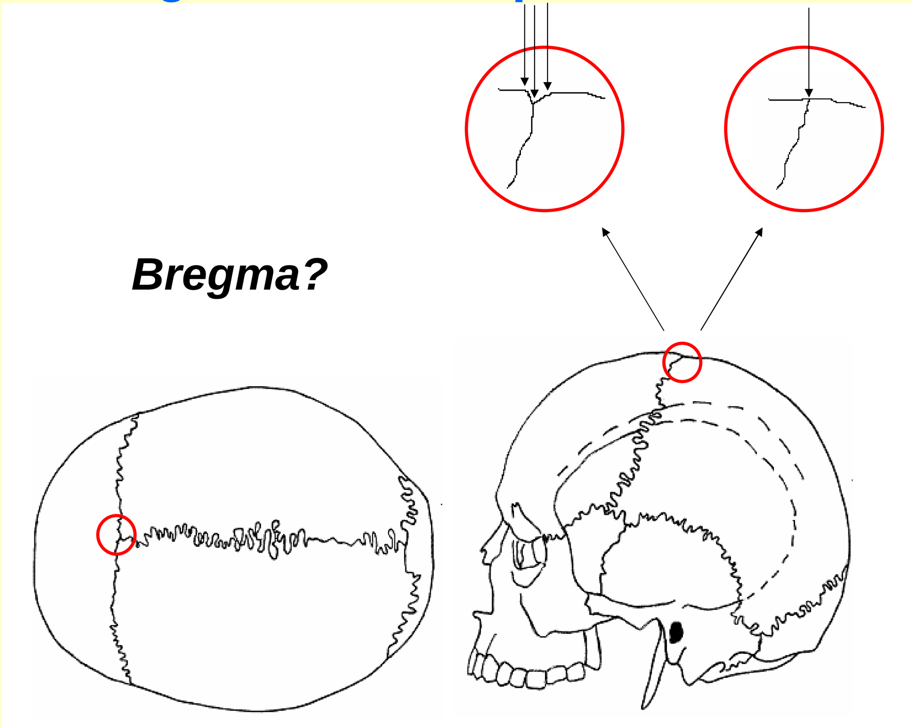

Figure 9. Defining landmark location to a very high precision may require more than a standard description.

References

[1]O'Higgins P, Jones N. 1998. Facial growth in *Cercocebus torquatus*: an application of three-dimensional geometric morphometric techniques to the study of morphological variation. J Anat 193:251-272.
[2]Lockwood CA, Lynch JM, Kimbel WH. 2002. Quantifying temporal bone morphology of great apes and humans; an approach using geometric morphometrics. J Anat 201:447-464.
[3]Chapman R. 1990. Conventional Procrustes approaches. In: Rohlf FJ, Bookstein FL, editors. Proceedings of the Michigan morphometrics workshop. Special publication No. 2. Ann Arbor: University of Michigan Museum of Zoology. p 251-267.
[4]Singleton M. 2002. Patterns of cranial shape variation in the Papionini (Primates: Cercopithecinae). J Hum Evol 42:547-578.
[5]Corner BD, Lele S, Richtsmeier JT. 1992. Measuring precision of three-dimensional landmark data. J Quan Anthropol 3:347-359.

Acknowledgements

We are indebted to S.J. Lycett for his perceptive observations throughout the design of these analyses and to R. P. Adams for advice on mathematical and statistical issues. Funding was gratefully received from the Cambridge Gates Trust and St Johns College, Cambridge (NvCT), the National Science Foundation, USA (BCF) and the Leverhulme Trust (MML).