AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 132:183–192 (2007)

Validation of Plaster Endocast Morphology Through 3D CT Image Analysis

P. Thomas Schoenemann,[1,2]* James Gee,[3] Brian Avants,[3] Ralph L. Holloway,[4] Janet Monge,[2,5] and Jason Lewis[6]

[1]*Department of Behavioral Sciences, University of Michigan-Dearborn, Dearborn, MI 48128*
[2]*Museum of Archaeology and Anthropology, University of Pennsylvania, Philadelphia, PA 19104*
[3]*Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104*
[4]*Department of Anthropology, Columbia University, New York, NY 10027*
[5]*Department of Anthropology, University of Pennsylvania, Philadelphia, PA 19104*
[6]*Department of Anthropological Sciences, Stanford University, Stanford, CA 94305*

KEY WORDS CT; computed tomography; endocast; validation

ABSTRACT A crucial component of research on brain evolution has been the comparison of fossil endocranial surfaces with modern human and primate endocrania. The latter have generally been obtained by creating endocasts out of rubber latex shells filled with plaster. The extent to which the method of production introduces errors in endocast replicas is unknown. We demonstrate a powerful method of comparing complex shapes in 3-dimensions (3D) that is broadly applicable to a wide range of paleoanthropological questions. Pairs of virtual endocasts (VEs) created from high-resolution CT scans of corresponding latex/plaster endocasts and their associated crania were rigidly registered (aligned) in 3D space for two *Homo sapiens* and two *Pan troglodytes* specimens. Distances between each cranial VE and its corresponding latex/plaster VE were then mapped on a voxel-by-voxel basis. The results show that between 79.7% and 91.0% of the voxels in the four latex/plaster VEs are within 2 mm of their corresponding cranial VEs surfaces. The average error is relatively small, and variation in the pattern of error across the surfaces appears to be generally random overall. However, inferior areas around the cranial base and the temporal poles were somewhat overestimated in both human and chimpanzee specimens, and the area overlaying Broca's area in humans was somewhat underestimated. This study gives an idea of the size of possible error inherent in latex/plaster endocasts, indicating the level of confidence we can have with studies relying on comparisons between them and, e.g., hominid fossil endocasts. Am J Phys Anthropol 132:183–192, 2007.

Reconstructing human brain evolution requires placing fossil endocranial evidence in the proper comparative context. Much of the comparative work that has been done to date has relied on endocasts made from modern human and ape samples (Radinsky, 1968; Holloway, 1975, 1976, 1978, 1981, 1983a; Falk, 1978; Holloway and de la Coste-Lareymondie, 1982; Falk et al., 1990). For the most part, these endocasts have been obtained through a fairly involved process resulting in rubber-coated plaster replicas (Radinsky, 1968; Holloway, 1975; Falk, 1978). Holloway (1975) describes the process as follows:

"Endocasts may also be made by applying liquid rubber latex to the inner cranial surface of a skull.... Successive layers are built up until a reasonable thickness, perhaps an eighth of an inch, is reached. The latex is cured by heat and then collapsed from the skull, either before or after stabilizing the dimensions with plaster."

While great care is taken in the production of these replicas, there are a number of steps in which a variety of subtle distortions could conceivably be introduced. Various foramina must be carefully filled (via the $\sim$3 cm diameter foramen magnum). Care must be taken to ensure that the liquid rubber latex covers the entire endocranial surface. The subsequently hardened thin rubber latex shell must then be collapsed from the skull and extracted through the foramen magnum (typically having anteroposterior dimensions approximately one-fifth the diameter of the latex shell, such that considerable deformation must occur to accomplish this). Finally, the extracted rubber latex shell is filled with plaster to stabilize it. This step is done under water, and since plaster initially poured in the latex shell is mostly water, the hydrostatic pressures should largely equalize. The plaster is heavier than the rubber latex shell, however, and thus has the potential to distort the final shape, especially if the rubber is not of uniform thickness as it conforms to the endocranial surface. In addition, plaster expands slightly when setting, producing potentially additional artifacts.

Although there have been some interesting qualitative studies of the extent to which the surface of the brain is actually reflected on the endocranial surface of the skull of the same individuals (Le Gros Clark et al., 1936), the extent to which the rubber-coated plaster replicas created by the above process typically match the actual

Grant sponsor: Penn University Research Fund; Grant sponsor: NSF Biological Resource Collections Award; Grant number: NSF 0447271.

*Correspondence to: Thomas Schoenemann, Department of Behavioral Sciences, University of Michigan-Dearborn, 4026 CASL Building, 4901 Evergreen Road, Dearborn, MI 48128, USA.
E-mail: ptoms@umd.umich.edu

Received 14 April 2006; accepted 8 August 2006

DOI 10.1002/ajpa.20499
Published online 13 November 2006 in Wiley InterScience (www.interscience.wiley.com).

endocranial surfaces has been difficult to determine with any confidence. Even if the endocranial surface could be exposed by sawing crania in half (or otherwise damaging the specimen), it is not clear how one could accurately measure the extent to which the surface of the plaster endocast actually matches the true endocranial surface using traditional techniques. However, analysis of high resolution 3-dimensional (3D) computed tomography (CT) scans of crania and their associated latex/plaster endocasts through the use of medical image analysis techniques allows for detailed comparisons between such pairs without damaging the crania.

The use of 3D CT to assess the morphology of both modern human and fossil hominid skeletal material has become increasingly common in recent years (Conroy and Vannier, 1985; Vannier et al., 1985; Conroy et al., 1998, 2000a,b; Ponce de Leon and Zollikofer, 1999; Recheis et al., 1999; Spoor et al., 2000; Tobias, 2001; Falk et al., 2005). CT scans are effectively sets of density measurements of each point (voxel) in a given 3D space. A CT scan of a cranium can be used to delimit its endocranial surface, thereby creating a virtual endocast (VE). A wide variety of medical imaging software packages are available to facilitate this process (see http://idoimaging.com/index.shtml for links to many free packages). Studies to date appear to generally have used a single global threshold value to identify the endocranial margin (the specific method of delineating the endocranial margin has not always been clearly identified in CT studies, however). With this method, all voxels with density values above this threshold are classified as containing bone. In most studies, some variant of a user-seeded region-growing algorithm is used on each successive slice of a 3D volume. This method groups all contiguous voxels containing density values within some range in the object of interest. The user indicates (places a "seed") where the algorithm should start, and also manually delimits areas where the algorithm should not be allowed to expand in to (as needed). For VEs, where the object of interest is the now-empty endocranial volume, all contiguous voxels with values below the user-selected threshold are included.

Because bone edges are marked by a gradient change of intensity over a distance of $\sim$1 mm (Weber et al., 1998; Spoor et al., 2000; Lewis et al., 2004), measurements can vary slightly depending on exactly where on this gradient the researcher chooses to define the edge (this is a function of the initial threshold chosen). Because the endocranial surface has varying thickness in different regions (very thick over the petrous portion of the temporal lobe, but very thin over the orbits), use of a global threshold for delimiting VEs is problematic. There is a built-in bias towards using a relatively low threshold, so as to avoid extensive manual delimitation (and consequent increased possibility of manual error) in areas where the bone is thin. The best solution is to use a method that focuses on the gradient change of density values over space. A manual method for simple linear measurements has been suggested by Spoor et al. (1993), but is completely impractical for high-resolution VEs which have $\sim$100,000 surface voxels. Some alternatives have been developed which typically use an automated expanding/contracting adaptive boundary (a "snake" or "balloon") to define boundaries (Spoor et al., 1993; Lobregt and Viergever, 1995; McInerney and Terzopoulos, 1995; Sethian, 1996; Yushkevich et al., 2005; Yushkevich et al., 2006). However, prior to the present study (see below) these have not been used to delimit VEs (nor have they been applied to paleoanthropological questions in general).

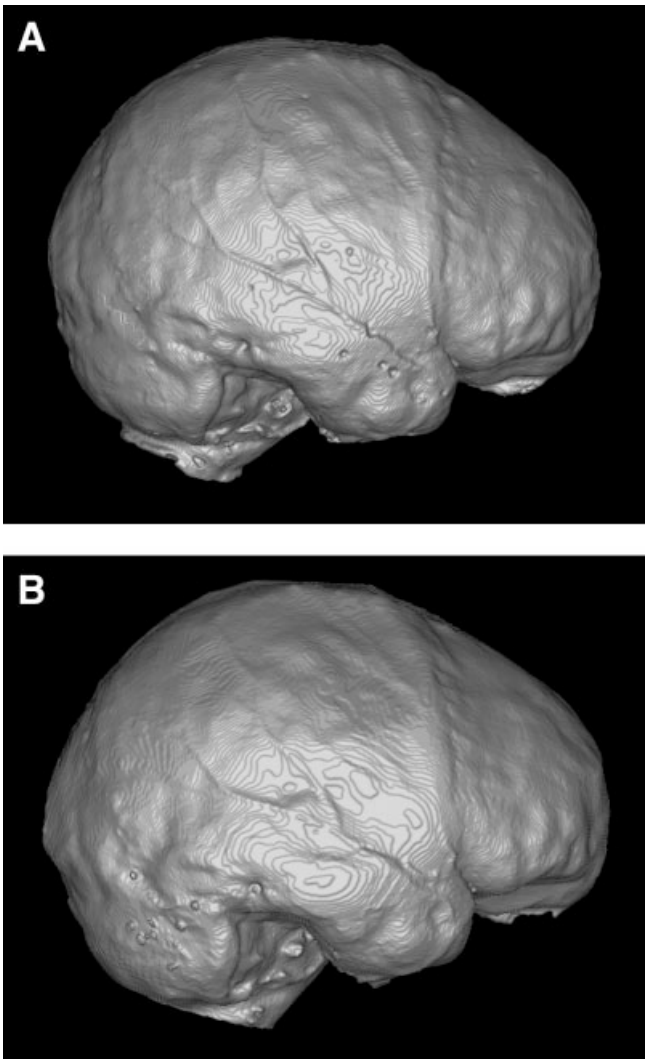

Fig. 1. 3D renderings of latex/plaster versus cranial VEs of specimen 512. (**A**) latex/plaster VE and (**B**) cranial VE.

Once delimited, VEs can be displayed on a computer screen as 3D renderings (e.g., Fig. 1), which can be oriented in any direction desired. In some studies the surface is translated into a virtual "wire mesh," which can be used to create a set of virtual tiles that approximate the delimited endocranial surface (Falk et al., 2005). The tiles smooth the surface to some extent, but are made small enough such that the important morphological features are not obscured. One advantage of such models is that the VE can be described in a more compact format (resulting in smaller file sizes), thereby decreasing computational resources required for rendering in 3D. For some types of mathematical analysis, however, VEs from the original volumetric dataset are more useful (e.g., for calculating local curvature over each point of the surface; see below).

In the present study, image segmentation and registration algorithms originally developed for medical image analysis research were applied to CT scans of four crania (two modern *Homo sapiens* and two *Pan troglodytes*)

along with associated latex/plaster endocasts of each crania (created by RLH) allowing for a detailed quantitative analysis of both extent and pattern of differences. We note that these techniques can be used to similarly compare any two objects of interest, such as fossil versus modern human crania, and will likely be of considerable use to a wide variety of paleoanthropological studies.

METHODS

Two human (*Homo sapiens*) crania (specimens 508 and 512, from the collection of RLH), and two chimpanzee (*Pan troglodytes*) crania (specimens L210ch13 and L212ch15 from the American Museum of Natural History), along with latex-coated plaster endocasts made by RLH from each of these specimens, were scanned on a General Electric Lightspeed CT scanner at the Hospital of the University of Pennsylvania. Specimens were scanned in coronal orientation while resting in a standard CT head cradle, with an exposure of 1782 ms at 120 kV tube voltage. The two human crania were scanned at a resolution of 0.44 mm $\times$ 0.44 mm $\times$ 1 mm (specimen 508) and 0.40 mm $\times$ 0.40 mm $\times$ 1 mm (specimen 512), and the two chimpanzee crania at 0.39 mm $\times$ 0.039 mm $\times$ 0.60 mm. The associated latex/plaster endocasts were scanned at a resolution of 0.36 mm $\times$ 0.36 mm $\times$ 0.60 mm (human specimens) or 0.28 mm $\times$ 0.28 mm $\times$ 0.50 mm (chimpanzee specimens). The resulting 16-bit images were transformed to 8-bit format with 0.75 mm (for the human crania) and 0.50 mm (for the chimpanzee crania) cubic voxels. This was done to decrease the file size (the algorithms applied require a very large amount of RAM memory; because chimpanzee endocranial volumes are so much smaller than the human endocranial volumes, finer resolution could be used without resulting in files too large for processing).

The endocranial surfaces of the crania were then delineated using a propagating balloon method that uses a gradient-based edge-detection algorithm, as implemented in SNAP (SNake Automatic Partitioning), a part of the Insight Toolkit (ITK) software tools package (www.itksnap.org, Yushkevich et al., 2005; Yushkevich et al., 2006). As discussed above, this method provides a more objective delineation of the edge than using a manually determined global threshold value to delineate the object of interest (Lewis et al., 2004; Lewis et al., in preparation). SNAP uses a propagating boundary ("snake" in 2D or "balloon" in 3D), which expands iteratively in all directions proportional to the gradients of the image intensities, such that the balloon can be induced to slow down and/or stop near edges or discontinuities of intensity. The images were smoothed before the balloon propagation step using a Gaussian filter (to decrease the likelihood that small high frequency edges in the image affect the overall segmentation). SNAPs algorithms have been validated against manual segmentation of the caudate nucleus of brain MRI, showing very high intra-rater reliabilities (intraclass $r \gg 0.95$) and correlated very strongly with manually segmented versions ($r \gg 0.96$, Yushkevich et al., 2005). The cranial VEs were delimited in the area of the foramen magnum in the following manner. A single plane was placed on a 3D view of the cranial VE (using a specialized "cutting" tool in SNAP) that most closely approximated the point where the inner (endocranial) and outer (ectocranial) layers of cortical bone meet on the foramen magnum margin. Because the opening of the foramen magnum is never perfectly flat, this plane was located so as to minimize the discrepancy in both the anteroposterior and lateral directions simultaneously. Our own work applying this method to endocranial volume assessment suggests that it is highly reliable (Lewis et al., in preparation). Volume estimates were then obtained by multiplying the number of voxels in each VE by the volume of each voxel. Lateral views of the 3D volume renderings of VEs produced by this technique are shown in Figure 1. Subjectively, the degree of similarity appears to be reasonably close. However, we quantified the degree of mismatch at each point on the surface with the following methods.

First, distance maps were calculated for the cranial VEs. Each voxel in the distance map volume contains a value equal to the distance that voxel is from the cranial VE surface (endocranial margin) in the direction normal to the closest surface point (Danielsson, 1980; Yoo et al., 2002). Distances extending outward from the surface are assigned positive values, while distances extending inward are assigned negative values. The surface itself has zero distance.

The cranial VEs were then rigidly registered (aligned) in 3D space to their corresponding latex/plaster VEs. In general, registration methods manipulate the pose of an image in a background space such that a cost function is minimized. In this case, the cost function is defined by mutual information (MI) (Wells et al., 1996), which has been shown to perform extremely well in many contexts and is robust to outliers (such as, in our data, the differences near the base of the skull, see below) and noise. MI performed well, in particular, in the Vanderbilt Retrospective Image Registration Evaluation Project (http://www.vuse.vanderbilt.edu/~image/registration/). MI measures the degree to which one can predict one image (in this case, the cranial VE) from knowledge about the other (the latex/plaster VE). If they are the same, one can perfectly predict the rotated image from the fixed image.

The goal of the algorithm is to search the space of rigid transformations until the MI is optimal. Our algorithm uses a minimal parameterization of the rigid transformation via a quaternion representation (Shoemake, 1985) and a multiresolution gradient descent optimization strategy to achieve near-optimal results. Four resolution levels (at 8, 4, 2, and 1 times the resolution of the input volume) were used, with a large number of iterations at each level to guarantee a large capture range and convergence. The algorithm starts with the coarsest resolution (eight times the input volume) and iterates up to 25,000 times until there is no change in its ability to predict the one image from the other (at that resolution). The algorithm is then iterated at the next higher resolution up to 25,000 times until there is no change in the prediction at that resolution, and so on for the successively higher levels of resolution. A final set of up to 25,000 iterations proceeds, again at the highest resolution, but this time with a finer degree of analysis of two volumes. The final transformation provides a rotation and translation that maps one image surface (the cranial VE) as closely as possible into the fixed image space (the latex/plaster VE).

This registration algorithm was applied not to the VEs themselves, but instead to curvature maps of the VEs, created by determining the local curvature at each point in the volume. These curvature maps were used because they contain more information regarding subtleties of

the 3D surface. Curvature was calculated on 3 voxel by 3 voxel local areas. Because the algorithm uses intensity gradients to calculate curvature (rather than surfaces per se), it calculates a curvature value for each point in the object of interest (i.e., each point with nonzero value). 3D curvature of a particular voxel is the average of the maximal and minimal 2D curvatures that are possible (given different planes oriented normal to the surface) at each point. 2D curvature is calculated as $1/r$, where r is the radius of a circle that matches the curvature of the line at that point. The algorithm for calculating curvature on 3D volumes is described in detail (Avants and Gee, 2003).

Once the cranial VEs were aligned with their corresponding latex/plaster VEs using the curvature maps, the transformation information was then used to rotate the corresponding distance maps for each cranial VE into alignment with their associated latex/plaster VEs. The intersection in 3D space of these rotated cranial VE distance maps with the surface voxels of their paired latex/plaster VEs results in detailed "difference-maps" of the distances between the corresponding surfaces of the paired VEs at each point (mapped onto the latex/plaster VE surfaces). Negative values on the resulting latex/plaster VE difference-maps therefore indicate areas where the latex/plaster VE surface is inside its corresponding cranial VE surface, positive values indicate where it is outside, and zero values where they exactly match. In total, there were 114,253 surface voxels (and hence, distance measurements) for the latex/plaster VE for human specimen 508, and 104,393 surface voxels for human specimen 512. The chimpanzee latex/plaster VEs had 93,434 voxels (for L210ch13) and 92,077 voxels (for L212ch15).

To determine the degree of error introduced by interpolation and/or registration errors of our method, we made six copies of a VE, rotated each one out of alignment in a different direction ($\pm 10^\circ$ in each of the X, Y, and Z dimensions), and then used our algorithms to align the original VE with each of these copies. If the method were perfect, distance maps between these pairs of identically shaped, but out of alignment VEs would result in zero values at each point. Actual distance maps for these pairs showed that the average distance across all surface voxels ranged from 0.0069 (SD = 0.2131) mm to 0.0089 (SD = 0.2157) mm for the six trials (mean average distance across the trials: 0.0078 mm). The RMSE (root mean squared error) across all voxels ranged from 0.2098 to 0.2159 mm (mean RMSE: 0.2129 mm). Across all trials, more than 99.7% of the voxels of the rotated versions were within 0.5 mm of the corresponding nonrotated copy. These results indicate that the size of the error introduced by rotation/registration/distance calculation is very small, once endocranial surfaces have been delimited.

All of the code used for the analyses in this study is available in ITK (www.itk.org), except for the surface curvature computation algorithm which, while written in ITK, is not currently part of the distribution. While it will be in the future, it is now available by contacting the authors.

To illustrate the extent of mismatch between latex/plaster and cranial VEs, histogram analyses of the difference-maps were then performed, resulting in graphs of the proportion of each latex/plaster CT VE surface that lays within a given distance of its corresponding cranial CT VE surface. The difference-maps were also color coded, allowing for a visual representation of the location and pattern of the areas of mismatch. This is particularly useful for understanding shape differences between any set of complex 3D shapes.

RESULTS

For human specimen 508, the volume of the cranial VE was 1659.5 ml, while the volume of the latex/plaster VE was 1642.0 ml. This represents a difference of −17.5 ml (latex/plaster VE smaller than cranial VE). For human specimen 512, the volume of the cranial VE was 1417.6 ml, while the volume of the latex/plaster VE was 1460.4 ml, a difference of +42.8 ml (latex/plaster VE larger than cranial VE). Chimpanzee specimen L210ch13 had a cranial VE volume of 334.7 ml, while its latex/plaster VE had a volume of 361.5 ml, a difference of +26.8 ml. Chimpanzee specimen L212ch15 had a cranial VE volume of 335.8 ml, while its latex/plaster VE had a volume of 347.6 ml, a difference of +11.8 ml. Thus, it appears that the latex/plaster method can result in absolute volume measurements that differ substantially from the CT VE method in terms of absolute milliliters, though for the human specimens these differences amount to only 1% (specimen 508) and 3% (specimen 512) of cranial VE values. For the chimpanzee specimens the absolute size of the error is similar to that found for the human specimens, but because of smaller volumes, the differences amount to 8% (L210ch13) and 4% (L212ch15). These differences are well within the range of variation of reported endocranial volume estimates in hominid specimens (De Miguel and Henneberg, 2001).

The average degree of mismatch at each point on the surfaces was found to be relatively small for all VE pairs. Figure 2 shows the histograms of the distances between the cranial and latex/plaster VEs for the human specimens (positive values indicate latex/plaster VE is outside of the corresponding cranial VE; negative values indicate the latex/plaster versions are inside). About 86.5% of the latex/plaster VE surface voxels for the specimen 508 comparison, and 91.0% of the voxels for the specimen 512 comparison are within 2-mm either way of their corresponding cranial VE surfaces. The root mean squared difference is 1.35 mm for specimen 508, and 1.43 mm for specimen 512 (the root mean squared difference is a better measure of overall fit, because negative and positive values tend to cancel each other out when calculating average distance). The chimpanzee specimens show very similar absolute values. Figure 3 shows histograms for the chimpanzee specimen VE pairs. About 79.7% of the latex/plaster VE surface voxels for L210ch13, and 86.5% of the voxels for L212ch15 are within 2-mm either way of their cranial VE surfaces. The root mean squared difference is 1.82 mm for specimen L210ch13, and 1.46 mm for specimen L212ch15. Thus, in absolute terms, VE pairs for both chimpanzee and human specimens are very similar.

The pattern of differences between the VE pairs for human and chimpanzee specimens is illustrated in Figures 4 and 5, respectively. In these images, green highlights areas where the two surfaces are within 0.5 mm of each other, red-to-yellow highlights the areas in which the latex/plaster VE surface is inside the cranial VE surface, and blue-to-white highlights the areas in which the latex/plaster VE surface is outside the cranial VE surface. The colored difference map for specimen 508 (Fig. 4a) clearly shows larger portions of the surface

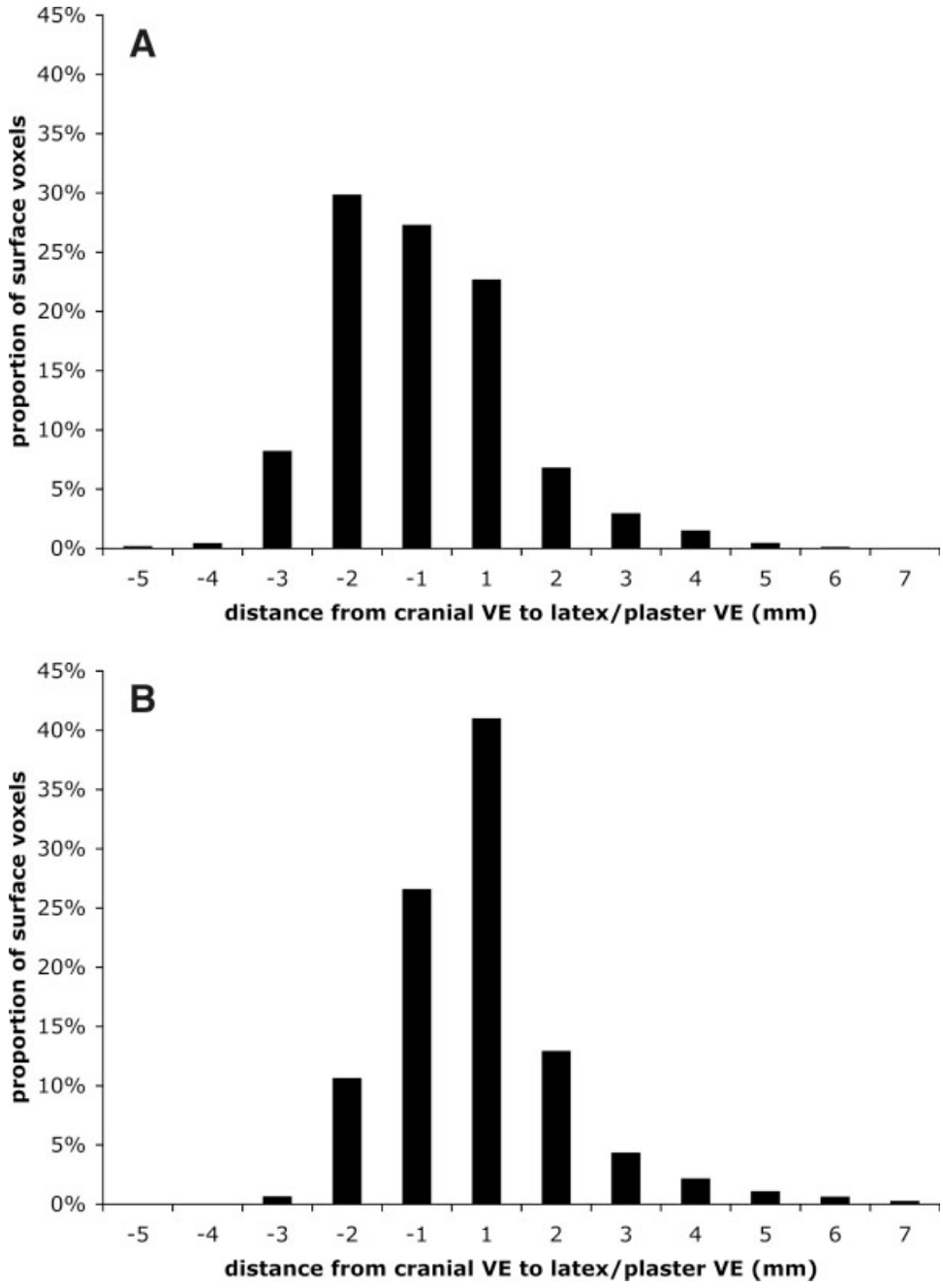

Fig. 2. Histograms of distances between human cranial and latex/plaster VEs. Negative values on the resulting latex/plaster VE difference-maps indicate areas where the latex/plaster surface is inside its corresponding cranial VE, and positive values indicate where it is outside. (**A**) Specimen 508 and (**B**) Specimen 512.

colored red–yellow, as compared to the map for specimen 512 (Fig. 4b). This is consistent with the finding that for specimen 508 the latex/plaster VE has a smaller volume than its corresponding cranial VE, while the reverse is true for specimen 512. Similarly, both chimpanzee specimens show larger portions colored blue–white, consistent with their latex/plaster VEs both having greater volumes than their corresponding cranial VEs.

These images show that foramen (particularly the foramen magnum) and the area around the cranial base generally, tend to show the largest differences between the cranial and latex/plaster VEs. All specimens show generally positive differences (blue–white: latex/plaster VE outside the cranial VE) in the area surrounding the sella turcica. The foramen magnum differs substantially in all specimens, with the latex/plaster VEs displaying a "lip" not present on the cranial VE. The delineation of the foramen magnum (and foramina generally) is somewhat arbitrary, as discussed below.

The lateral, anterior, posterior, and superior aspects of the endocranial surface do not have the same issues with respect to arbitrary boundary designation as seen in the inferior cranial base areas. In general, the results suggest: 1) the differences for lateral, anterior, posterior, and superior areas are on the whole smaller than those found in the inferior regions, but 2) the pattern of differences in different specimens clearly varies. In general, there does not appear to be a consistent pattern of mismatch across the specimens in this study, suggesting that errors introduced by this method tend to be random

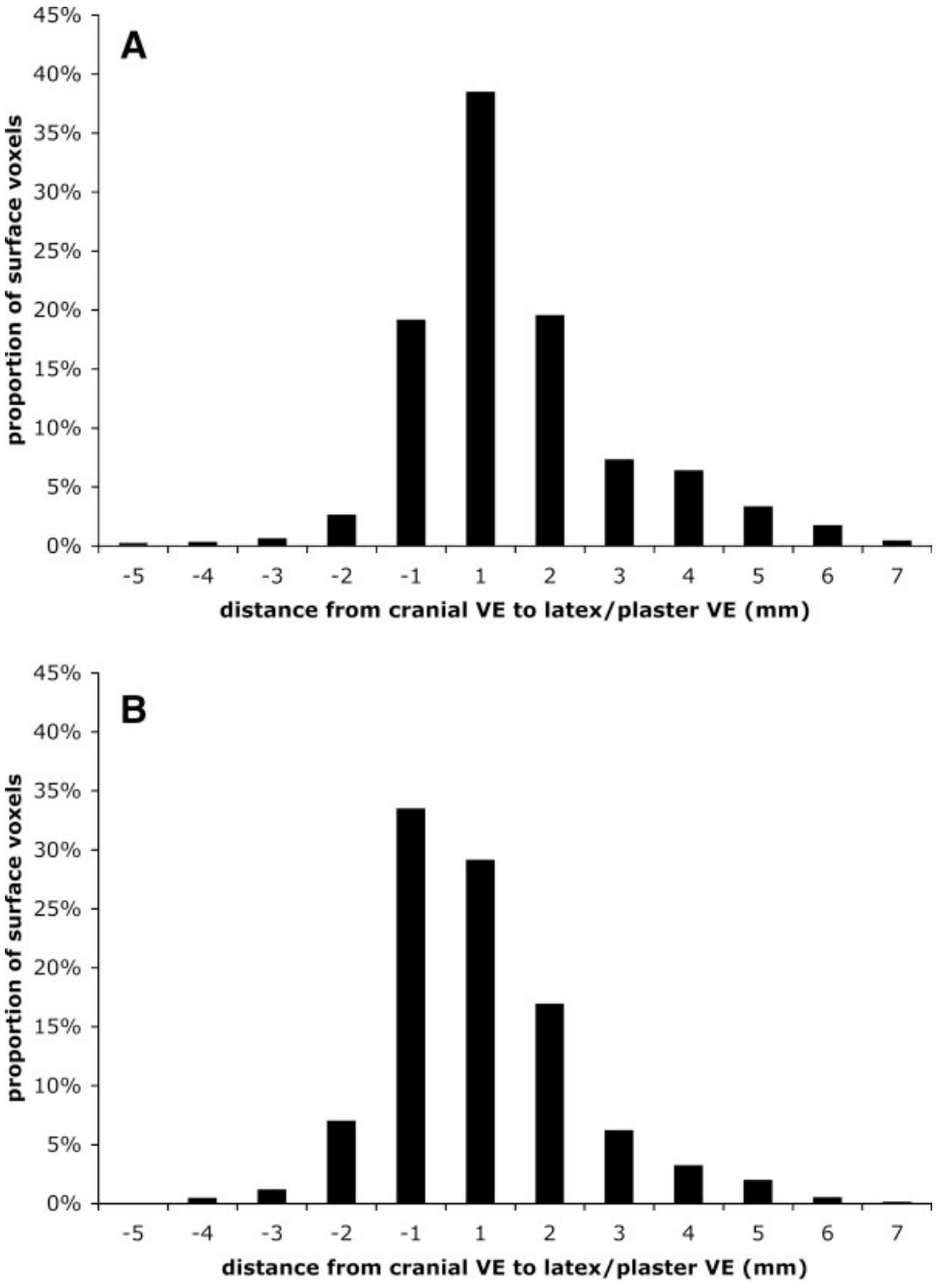

Fig. 3. Histograms of distances between chimpanzee cranial and latex/plaster VEs (see FIGURE 2 text for details). (**A**) Specimen L210ch13 and (**B**) Specimen L212ch15.

(rather than systematic). To the extent that these latex/plaster endocasts are representative of the method generally, this study suggests that relatively small errors are introduced.

A few areas hint at the possibility of nonrandom error. In both human specimens, and for both hemispheres, portions of the latex/plaster VEs in the area of Broca's cap (which approximates the endocranial surface over Broca's area of the brain; see below) protrude less (and hence are colored red–orange) than the cranial VEs. The same areas on the chimpanzee specimens tend in the opposite direction, with the left hemisphere of specimen L210CH13 and the right hemisphere of specimen L212CH15 showing significant portions in the area surrounding Broca's cap that protrude more (colored blue) on the latex/plaster VEs, and the other hemispheres for the most part suggesting little difference between the latex/plaster versus cranial VEs (mostly colored green, with a few red portions). The differences in the human comparisons are on the order of $\sim$1–2 mm.

The maps in Figures 4 and 5 also show that the temporal poles are generally overestimated on the latex/plaster endocasts. The left temporal poles of both human specimens are color-coded blue, though for specimen 508 only the left temporal pole appears to be substantially blue. For the chimpanzee specimens, both temporal poles for L210ch13 and the right temporal pole for L212ch15 are blue–white.

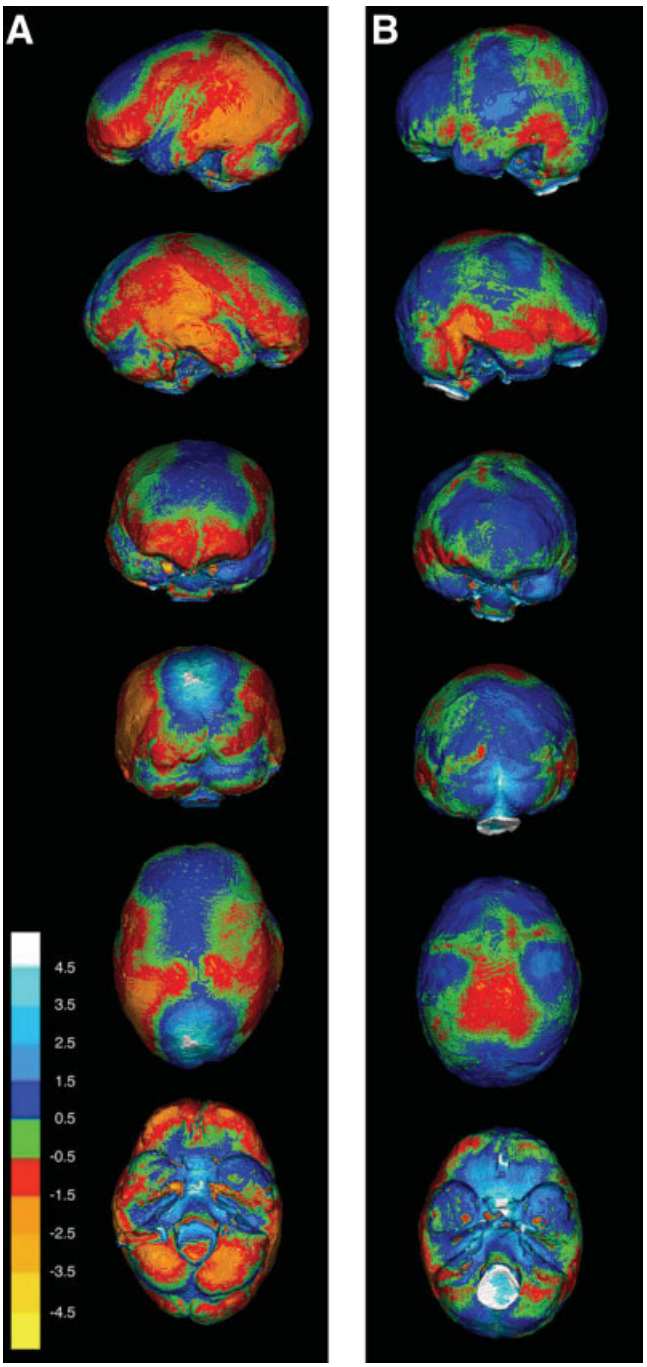

Fig. 4. Patterns of mismatch between human cranial and latex/plaster VEs. Latex/plaster VEs are color-coded to highlight the degree of mismatch with their corresponding cranial VEs at each point. Green highlights areas where the two surfaces are within 0.5 mm of each other, red-to-yellow highlights the areas in which the latex/plaster VE surface is inside the cranial VE surface, and blue-to-white highlights the areas in which the latex/plaster VE surface is outside the cranial VE surface. The views depicted are, from top to bottom: left lateral, right lateral, anterior, posterior, superior (anterior facing up), inferior (anterior facing up). (**A**) Specimen 508, with color code distance key (numbers in mm) and (**B**) Specimen 512 (color-coding as in a).

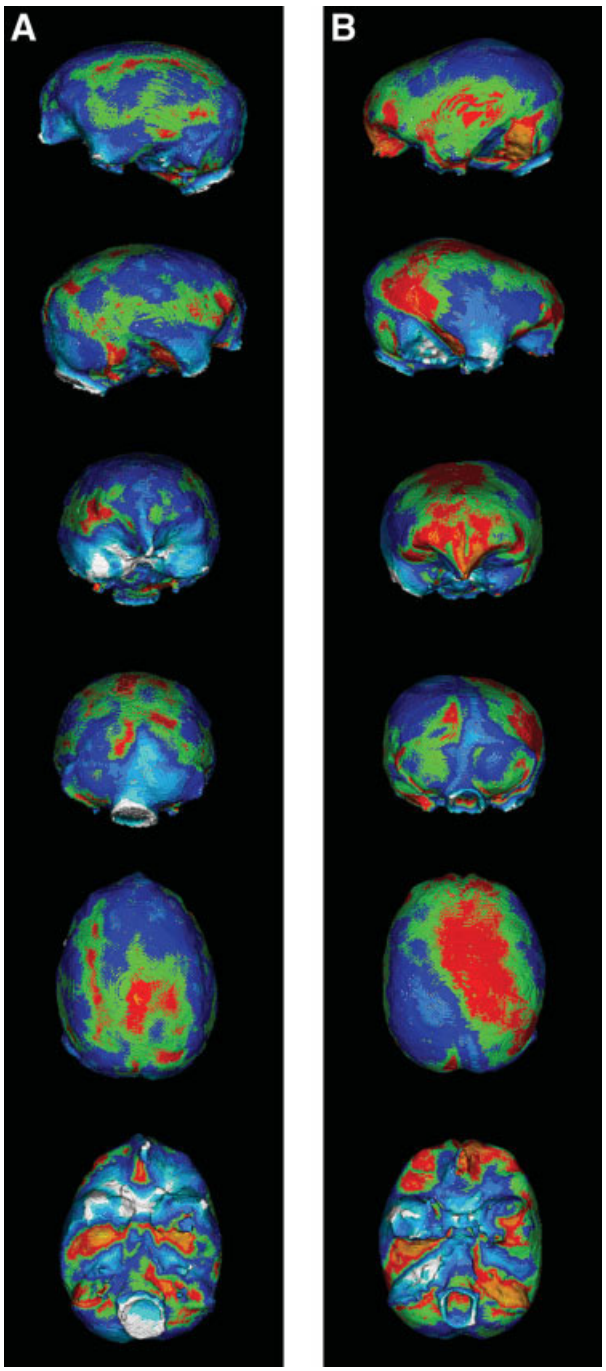

Fig. 5. Patterns of mismatch between chimpanzee cranial and latex/plaster VEs. Latex/plaster VEs are color-coded to highlight the degree of mismatch with their corresponding cranial VEs at each point (see Fig. 4 text for details). The views depicted, their arrangement, and color-coding are as in Figure 4. (**A**) Specimen L210CH13 and (**B**) Specimen L212CH15.

DISCUSSION

Overall, methodological errors involved in latex/plaster endocast creation appear to be relatively small, with an RMSE of less than 2 mm overall. As one might expect, the pattern of error is largely idiosyncratic, though judging from these specimens the foramen magnum, sella turcica, temporal poles, and (in humans at least) the area overlaying Broca's area, show somewhat similar patterns of error. The small sample size of the present study limits our ability to make generalizations, but these areas hint at the possibility of nonrandom bias.

The fact that the area of greatest mismatch is in the cranial base is not unexpected, given that most of the foramina are located there. Foramina present inevitable difficulties for segmentation and comparison, because the exact point where the endocranial volume ends and a foramen starts is inherently arbitrary. The differences highlighted in the region of the foramen magnum in the present study are a reflection of the fact that the latex/plaster VEs "spill out" in this area, whereas the cranial VEs were specifically segmented to avoid this. This difference is not particularly troublesome, however, because the arbitrary nature of the delineation of the foramen magnum has been implicitly recognized by researchers, and is not the focus of paleoneurological interest. However, it is likely to contribute to variation in estimates of endocranial volumes in fossil specimens, whether or not they are latex/plaster or CT-derived VEs (e.g., the LB1 endocast shown in Falk et al., 2005 has a distinct foramen magnum "lip"). There does not appear to be a clear discussion in the literature of exactly how this area is (or should be) delimited.

That the cranial base region generally shows the largest positive deviations of the latex/plaster VEs from the cranial VEs is partly attributable to the fact that this portion is the last to be filled with plaster. In addition, plasticine modifications must be made to the cranium, prior to pouring of latex, so that the subsequent removal through the foramen magnum does minimal damage to the cranial base. The sella tursica is sometimes filled with plasticine to avoid damage when retracting the rubber endocast. How much each of these aspects contributes to distortion is unclear. However, the cranial base portions of endocasts have not been the focus of interest by researchers, and this study illustrates that some caution would need to be taken in interpretations of this area based on latex/plaster endocasts.

Though highly tentative given the sample size, the similar degrees of underestimation found in the region of Broca's cap in both human specimens, and the tendency for the chimpanzee specimens in the opposite direction, may be of interest for future studies. The specific neurobehavioral significance of Broca's cap is unknown, but given its location it is assumed to be potentially relevant to the evolution of language (Sherwood et al., 2003), and thus has been the focus of significant interest in paleoneurology (Tobias, 1975; Falk, 1983; Holloway, 1983b; Broadfield et al., 2001).

The pattern of overestimation in temporal lobe regions evident in the present study may also be of future interest. The degree of difference is again relatively small. Studies of actual brains suggest that human temporal lobes are larger than predicted based on ape scaling trends (Rilling and Seligman, 2002). The temporal lobes are known to play key roles in memory, emotion, and language (Carpenter and Sutin, 1983; Damasio and Damasio, 1992), and are thus of particular interest. Potential overestimation of the temporal poles by latex/plaster endocasts should be addressed in any future studies that rely on them.

Future studies with significantly larger sample sizes will be needed before any definitive conclusions could be made about possible systematic bias of latex/plaster endocasts in either the Broca's cap or temporal lobe regions. There does not appear to be a simple way to significantly speed up the analyses by limiting computations to a specific area of the endocranial surface, since the entire volume needs to be registered (aligned) before any localized measurements are made (the registration step is by far the most computationally intensive). However, refinements of the software code and increasing computer processor speed will likely help future studies. In any case, the present study does suggest that errors of $\sim$1–2 mm are not uncommon in different regions of the endocranial surface.

In addition to the specific findings of this study, we wish to point out that the method of direct, point-by-point comparison between endocasts outlined here can be easily and directly generalized to the comparison of any two objects of interest, specifically of interest to hominid paleontology and comparative anatomy: fossil-to-modern, fossil-to-fossil, and comparative studies of modern species comparisons. The method only requires 3D representations of the objects of interest, which can be obtained through CT, laser scanning, MRI, etc. This allows for a global 3D high-resolution analysis of the differences at each point. Furthermore, it is technically possible to process a population of objects (e.g., modern human crania, ape articular surfaces, etc.) to create an atlas describing the average shape (as well as the variation at each point on that shape) across the sample. In addition, nonrigid (morphing) methods can be applied to such images, allowing for the quantification of the degree of localized distortion necessary to map one object (e.g., endocast, cranium, fossil fragment) into another (e.g., an atlas representing the average shape of some population). These techniques allow for rigorous quantitative analyses of shape differences between fossil and modern forms (endocrania, femur condyles, vertebrae, innominates, etc.) and are already being applied to a variety of questions of evolutionary interest (e.g., MacLeod et al., 2003; Van Essen, 2005; Zilles, 2005; Avants et al., 2006). We have begun the initial steps of creating such atlases for the endocranial form for both modern *Homo sapiens* and *Pan troglodytes* specimens from the Open Research Scan Archive (http://grape.anthro.upenn.edu/$\sim$lab/pennct/; scans are freely accessible to qualified researchers for a nominal fee (Monge et al., 2004)). The creation of atlases of endocranial form, and the methods outlined in the present study, will likely be of considerable importance for quantitative analyses of fossil specimens in the future.

ACKNOWLEDGMENTS

We thank two anonymous reviewers and Clark Spencer Larsen for comments and suggestions that improved an earlier draft of this manuscript. We thank Felicia Jefferson for taking the CT images used in this study, and Nick Bryan and the Department of Radiology, University of Pennsylvania Health System, for access to the scanners. We also thank Ian Tattersall and Kenneth Mowbray from the American Museum of Natural History

for allowing us to scan (and make generally available) the *Pan troglodytes* crania used in this study. Daniel Glotzer and Ariel Singer helped with scanning the endocasts. We also thank the University of Pennsylvania Museum of Archaeology and Anthropology, the University of Pennsylvania, and NSF for supporting the Open Research Scan Archive. This research benefited from the use of the Insight Segmentation and Registration Toolkit (ITK), an open source software developed as an initiative of the U.S. National Library of Medicine and available at www.itk.org. The multiplatform configuration tool CMake was used for configuring ITK and facilitating its use for our project. CMake was partially funded by the U.S. National Library of Medicine as part of the Insight Toolkit project. CMake is an open source system and is freely available at www.cmake.org.

LITERATURE CITED

Avants B, Gee J. 2003. The shape operator for differential analysis of images. Inf Process Med Imaging 18:101–113.

Avants BB, Schoenemann PT, Gee JC. 2006. Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex. Med Image Anal 10:397–412.

Broadfield DC, Holloway RL, Mowbray K, Silvers A, Yuan MS, Márquez S. 2001. Endocast of Sambungmacan 3 (Sm 3): a new *Homo erectus* from Indonesia. Anat Rec 262:369–379.

Carpenter MB, Sutin J. 1983. Human neuroanatomy. Baltimore: Williams & Wilkins.

Conroy GC, Falk D, Guyer J, Weber GW, Seidler H, Recheis W. 2000a. Endocranial capacity in STS 71 (*Australopithecus africanus*) by three-dimensional computed tomography. Anat Rec 258:391–396.

Conroy GC, Vannier MW. 1985. Endocranial volume determination of matrix-filled fossil skulls using high-resolution computed tomography. In: Tobias P, editor. Hominid evolution: past, present, and future. New York: Alan R Liss. p 419–426.

Conroy GC, Weber GW, Seidler H, Recheis W, Zur Nedden D, Mariam JH. 2000b. Endocranial capacity of the Bodo cranium determined from three-dimensional computed tomography. Am J Phys Anthropol 113:111–118.

Conroy GC, Weber GW, Seidler H, Tobias PV, Kane A, Brunsden B. 1998. Endocranial capacity in an early hominid cranium from Sterkfontein, South Africa. Science 280:1730–1731.

Damasio AR, Damasio H. 1992. Brain and language. Scientific American 267(3):89–95.

Danielsson P-E. 1980. Euclidean distance mapping. Comput Graph Image Process 14:227–248.

De Miguel C, Henneberg M. 2001. Variation in hominid brain size: how much is due to method? Homo 52:3–58.

Falk D. 1978. Brain evolution in old world monkeys. Am J Phys Anthropol 48:315–319.

Falk D. 1983. Cerebral cortices of East African early hominids. Science 221:1072–1074.

Falk D, Hildebolt C, Cheverud J, Vannier M, Helmkamp RC, Konigsberg L. 1990. Cortical asymmetries in frontal lobes of Rhesus monkeys (*Macaca mulatta*). Brain Res 512: 40–45.

Falk D, Hildebolt C, Smith K, Morwood MJ, Sutikna T, Brown P, Jatmiko, Saptomo EW, Brunsden B, Prior F. 2005. The brain of LB1, *Homo floresiensis*. Science 308:242–245.

Holloway RL. 1975. The role of human social behavior in the evolution of the brain. New York: American Museum of Natural History.

Holloway RL. 1976. Some problems of hominid braincast reconstruction, allometry, and neural reorganization. In: Colloquium VI of the IX Congress of the UISPP, Nice, 1976. p 69–119.

Holloway RL. 1978. The relevance of endocasts for studying primate brain evolution. In: Noback CR, editor. Sensory systems in primates. New York: Plenum. p 181–200.

Holloway RL. 1981. Volumetric and asymmetry determinations on recent hominid endocasts: Spy I and II, Djebel Ihroud I, and the Sale *Homo erectus* specimens, with some notes on Neanderthal brain size. Am J Phys Anthropol 55: 385–393.

Holloway RL. 1983a. Cerebral brain endocast pattern of *Australopithecus afarensis* hominid. Nature 303:420–422.

Holloway RL. 1983b. Human paleontological evidence relevant to language behavior. Hum Neurobiol 2:105–114.

Holloway RL, de la Coste-Lareymondie MC. 1982. Brain endocast asymmetry in pongids and hominids: some preliminary findings on the paleontology of cerebral dominance. Am J Phys Anthropol 58:101–110.

Le Gros Clark WE, Cooper DM, Zuckerman S. 1936. The endocranial cast of the chimpanzee. J R Anthropol Inst Great Brit Ireland 66:249–268.

Lewis JE, Schoenemann PT, Monge J. 2004. Endocranial capacity estimated from 3-D CT: methodological issues. Am J Phys Anthropol 123 (Suppl 38):135.

Lobregt S, Viergever MA. 1995. A discrete dynamic contour model. IEEE Trans Med Imaging 14:12–24.

MacLeod CE, Falk D, Mohlberg H, Zilles K. 2003. Patterns of surface shape in great ape endocasts. Am J Phys Anthropol 120 (Suppl 36):143–144.

McInerney T, Terzopoulos D. 1995. Medical image segmentation using topologically adaptable snakes. In: Ayache N, editor. Computer vision, virtual reality and robotics in medicine. Berlin: Springer-Verlag. p 92–101.

Monge J, Schoenemann PT, Lewis JE, Glotzer D. 2004. The CT database at the University of Pennsylvania Museum. Am J Phys Anthropol 123 (Suppl 38):149.

Ponce de Leon MS, Zollikofer CPE. 1999. New evidence for Le Moustier 1: computer-assisted reconstruction and morphometry of the skull. Anat Rec 254:474–489.

Radinsky LB. 1968. A new approach to mammalian cranial analysis, illustrated by examples of prosimian primates. J Morphol 124:167–180.

Recheis W, Macchiarelli R, Seidler H, Weaver DS, Schafer K, Bondioli L, Weber GW, zur Nedden D. 1999. Re-evaluation of the endocranial volume of the Guattari 1 Neandertal specimen (Monte Circeo). Coll Antropol 23:397–405.

Rilling JK, Seligman RA. 2002. A quantitative morphometric comparative analysis of the primate temporal lobe. J Hum Evol 42:505–533.

Sethian JA. 1996. Level set methods; evolving interfaces in geometry, fluid mechanics, computer vision, and materials science. Cambridge: Cambridge University Press.

Sherwood CC, Broadfield DC, Holloway RL, Gannon PJ, Hof PR. 2003. Variability of Broca's area homologue in African great apes: implications for language evolution. Anat Rec 271A(2):276–285.

Shoemake K. 1985. Animating rotation with quaternion curves. Comput Graph 19:245–254.

Spoor CF, Zonneveld FW, Macho GA. 1993. Linear measurements of cortical bone and dental enamel by computed tomography: applications and problems. Am J Phys Anthropol 91: 469–484.

Spoor F, Jeffery N, Zonneveld F. 2000. Using diagnostic radiology in human evolutionary studies. J Anat 197:61–76.

Tobias PV. 1975. Brain evolution in the Hominoidea. In: Tuttle RH, editor. Primate functional morphology and evolution. The Hague: Mouton. p 353–392.

Tobias PV. 2001. Re-creating ancient hominid virtual endocasts by CT-scanning. Clin Anat 14:134–141.

Van Essen DC. 2005. Surface-based comparisons of macaque and human cortical organization. In: Dehaene S, Duhamel J-R, Hauser MD, and Rizzolatti G, editors. From monkey brain to human brain. Cambridge, Massachusetts: MIT Press. p 3–19.

Vannier MW, Conroy GC, Marsh JL, Knapp RH. 1985. Three-dimensional cranial surface reconstructions using high-resolution computed tomography. Am J Phys Anthropol 67: 299–311.

Weber GW, Recheis W, Scholze T, Seidler H. 1998. Virtual anthropology (VA): Methodological aspects of linear and volume measurements—first results. Coll Antropol 22:575–584.

Wells WM, 3rd, Viola P, Atsumi H, Nakajima S, Kikinis R. 1996. Multi-modal volume registration by maximization of mutual information. Med Image Anal 1:35–51.

Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S, Metaxes D, Whitaker R. 2002. Engineering and algorithm design for an image processing API: a technical report on ITK—The insight toolkit. In: Westwood J, editor. Proceedings of medicine meets virtual reality. Amsterdam: IOS Press. p 586–592.

Yushkevich PA, Piven J, Hazlett HC, Ho S, Gee JC, Gerig G. 2005. User-guided level set segmentation of anatomical structures with ITK-SNAP. Insight Journal 1 (Special Issue on ISC/NA-MIC/MICCAI Workshop on Open-Source Software).

Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, and Gerig G. 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31(3):1116–1128.

Zilles K. 2005. Evolution of the human brain and comparative cyto- and receptor architecture. In: Dehaene S, Duhamel J-R, Hauser MD, Rizzolatti G, editors. From monkey brain to human brain. Cambridge, Massachusetts: MIT Press. p 41–56.