5-1-2010

Taxonomic implications of basicranial variation in Australopithecus africanus

Timothy R. Petersen

Timothy R. Petersen
Candidate

Anthropology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

, Chairperson

Brian C. Richmond

TAXONOMIC IMPLICATIONS OF BASICRANIAL VARIATION IN AUSTRALOPITHECUS AFRICANUS

BY

TIMOTHY R. PETERSEN

Bachelor of Arts, Psychology, Rice University, 1994
Master of Arts, Anthropology, University of Houston, 2000

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Anthropology

The University of New Mexico
Albuquerque, New Mexico

May, 2010

To my family

Acknowledgements

I owe thanks to many people. Were it not for their patience, support, and encouragement, this work probably would not exist, and definitely would not have seen the light of day. Here are some of the more notable, in no particular order.

This work was funded by a Doctoral Dissertation Improvement Grant from the National Science Foundation, grant no. BCS-0451969.

The staff of the museums I visited were always welcoming and helpful: the Transvaal Museum (Northern Flagship Institution), the Anatomy Department at the University of the Witwatersrand, the Powell-Cotton Museum, le Musée Royal de l'Afrique Centrale, the Peabody Museum at Harvard University, and the Cleveland Museum of Natural History. Stephany Potze at the Transvaal deserves special commendation for her hospitality.

Thanks to Susan McIntosh for showing me how interesting osteology could be, and for the best introduction to a Human Evolution course I've ever seen; and to Rebecca Storey and the other UH faculty for showing me how to do social sciences research.

Thanks to Ozzie Pearson for welcoming me to UNM in the first place, and then once again when I decided that his line of research was much more interesting than I would originally admit. He taught me more than would fit in some mere book, and the simple fact that I was his student meant that I was well-received on three continents. More people should be as lucky as I was with my advisor.

Thanks to Brian Richmond for being a part of this effort, and for taking the first step toward disseminating the work herein to a broader audience. His help has already started to pay off.

Sherry Nelson provided insight into some interesting aspects of the data I have accumulated. She showed me avenues that I hadn't realized were there, and because of her, I'll try never to forget that there are always several productive angles.

To the extent that this work is clear, says what I meant it to say, and makes arguments that are halfway cogent, I have to thank Keith Hunley for his editing efforts. The benefit of having a good editor cannot be overstated.

Jane Buikstra provided invaluable guidance and support to me as a young graduate student, and as I undertook this project. It would not have been what it is now, much less funded and underway, without her assistance.

The late Charles Lockwood also provided early guidance and encouragement. He will be missed.

Thanks to my friends for making graduate school tolerable.

Thanks to my family (including the in-laws!) for their unwavering support and patience. They have put up with a lot, and never held it against me.

And thanks most of all to my wife. Words fail me here, so I will just take one more opportunity to say that she means everything to me.

TAXONOMIC IMPLICATIONS OF BASICRANIAL VARIATION IN AUSTRALOPITHECUS AFRICANUS

BY

TIMOTHY R. PETERSEN

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Anthropology

The University of New Mexico
Albuquerque, New Mexico

May 2010

TAXONOMIC IMPLICATIONS OF BASICRANIAL VARIATION IN AUSTRALOPITHECUS AFRICANUS

by

Timothy R. Petersen

Bachelor of Arts, Rice University, 1994
Master of Arts, University of Houston, 2000

Doctor of Philosophy, Anthropology
University of New Mexico, 2010

ABSTRACT

Although it was discovered 85 years ago, *Australopithecus africanus* remains a source of contention among paleoanthropologists. Uncertainty about the fossils' taxonomic unity has resulted in controversy about their place in hominin phylogeny. This work addresses their taxonomy through application of three-dimensional morphometrics followed by analysis of their patterns of variation in traditional morphological characters. This sequential approach lends more support to the conclusions than would either technique alone. The cranial base was selected as the focus of the analyses because it preserves well and is likely to capture taxonomically-important variation. This inference is supported by the finding herein that the cranial bases of *Pan troglodytes* and *Pan paniscus* are morphometrically distinct. Morphometrically, the sample of basicranial specimens typically assigned to *A. africanus*

is shown here to be slightly more variable than is a broad pooled-species sample of *Pan* crania. The fossils' pairwise Procrustes distances were compared to the distribution of similar distances within that *Pan* sample, and their percentile scores in its distribution were ordinated to reveal underlying patterns in their variation. This indirect approach to the comparisons permitted more fossils to be analyzed together than would otherwise have been possible. In these comparisons, several *A. africanus* specimens are shown to have distinct morphometric shapes in their basicrania. Of these, Sts 19, Sts 25, and Stw 580 also have distinct patterns of traditional morphological characteristics. The consistency with which these three specimens are distinguished indicates that they are likely to belong to a taxon other than *A. africanus*, but the fragmentary nature of Stw 580 renders this conclusion more tentative for that particular specimen. These findings do not support the hypothesis that the Sterkfontein and Makapansgat fossil assemblages represent different taxa, nor the hypothesis that the Sterkfontein sample contains approximately equal numbers of specimens from two distinct species.

Contents

List of Figures

List of Tables

Chapter 1. Introduction

Raymond Dart's (1925) announcement that the species represented by a small-brained juvenile South African fossil was not only a bipedal ape but probably ancestral to humans was not universally accepted, and once was publicly called "preposterous" (Keith, 1925b:11; 1925a). His colleague Robert Broom's use of the Taung specimen and later ones from Sterkfontein and Makapansgat to refute claims that the Piltdown material represented direct human ancestors helped to perpetuate the controversy (Strkalj et al., 2005). One of Broom's students, Phillip Tobias (2001; 1985), has reviewed the history of scientific views of *Australopithecus*, and notes that general acceptance did not occur until the 1950s. He argues that a major stimulus to this acceptance was the "conversion" (2001:17) of two of Dart and Broom's erstwhile opponents, Sirs Arthur Keith (1947) and Wilfrid le Gros Clark (1946), upon reading the large postwar monograph detailing the fossils' morphology (Broom and Schepers, 1946). Since that time, many more African fossil hominin taxa have been identified, and the ancestor/descendant relationships among many of them remain matters of contention. Although it is now accepted as a hominin, *A. africanus* remains a periodic focus of taxonomic disagreement. This project is designed to address the question of whether the taxon *Australopithecus africanus* provides a useful way to describe the assigned specimens, or alternatively whether variation among the Sterkfontein and Makapansgat hominin fossils would be better conceptualized if the taxon were split.

The status of Australopithecus africanus

The most commonly accepted taxonomy for *A. africanus* fossils is Robinson's (1954). As it gained ascendancy, two other prior arguments fell into broad, but not unanimous, disfavor. One was that specimens from different sites (then Sterkfontein, Makapansgat, and Taung) constituted different species or even different genera. Clarke (1994; 2008) has briefly described the initial history of the taxonomy of these fossils following Taung. Broom (1936) first referred the TM 1511 Sterkfontein specimen to *Australopithecus transvaalensis* (i.e. the same genus but a different species than the Taung specimen). Two years later (1938), however, when a large-toothed juvenile specimen (mandible fragment TM 1516 and canine Sts 50) was found, he transferred the entirety of the Sterkfontein material to a new genus (*Plesianthropus*). The discovery of a third site, Makapansgat, resulted in a third species when Dart assigned its fossils to *A. prometheus* (Dart, 1948a, b). The other hypothesis was that these fossils and the consistently large-toothed and large-faced *A. robustus*[1] represented two sexes as opposed to two species, but the later discovery that Swartkrans postdated Sterkfontein and Makapansgat rendered this proposal much less tenable (reviewed in White et al., 1981; and Wood and Richmond, 2000). Robinson's (1954) division of the South African australopithecines into gracile (*A. africanus*) and robust (*A. robustus*) species remains the most common interpretation.

[1] The generic nomen *Australopithecus* is retained here for *A. afarensis* and the "robust" group of hominin taxa, but is not intended to imply support or rejection of any particular phylogenetic hypothesis. This retention is provisional, in the expectation that *Australopithecus* will eventually be split because it is paraphyletic (Footnote 1 of Grine et al., 2006; Strait and Grine, 2004; Wood and Richmond, 2000; but see Leakey et al., 2001). The details of a new taxonomy, however, have yet to be agreed upon, despite identification of this problem over twenty years ago (Delson, 1987). Any changes to the accepted alpha taxonomy of *A. africanus* due to this and/or other work would likely stimulate a revision of these species' phylogeny and therefore genus-level taxonomy.

Despite the general agreement following Robinson (1954), the taxonomy of specimens assigned to *A. africanus* has again become controversial. Particular specimens such as Stw 53, Sts 19, Stw 505, Sts 71, Stw 252, and even Taung itself[2] have come under periodic scrutiny as possibly belonging to *Homo*, the robust australopithecine clade, or another novel taxon. Numerous authors (e.g. Clarke, 2008; Plavcan, 2003; Lockwood and Tobias, 2002; Strait et al., 1997; Lockwood, 1997; Calcagno et al., 1997; Clarke, 1994; Falk, 1990; Picq, 1990) have expressed concern about the taxonomic implications of cranial and/or dental variation among the specimens assigned to *A. africanus*, but others remain confident that they represent a single taxon (e.g. McNulty et al., 2006; McCollum, 2000; Ahern, 1998; Wood, 1994; Tobias, 1989; White et al., 1981). Few, however, continue to argue that specimens outside South Africa belong to *A. africanus*, i.e. that it contains too little variation (e.g. Robinson, 1954; Wolpoff, 1971; Howell, 1978; Tobias, 1980; Holloway, 1981). As discussed below, new discoveries have lent more urgency to this question.

Dating controversies

This project contributes to resolution of the taxonomic controversy through analysis of basicranial variation among specimens assigned to *Australopithecus africanus*, which existed between approximately 3.0 and 2.5 million years ago (Ma) (Vrba, 1982; Partridge, 1986; Delson, 1988; McKee, 1993; Schwarcz et al., 1994; also summarized in Klein, 2009). These dates, which are central to taxonomic and other

[2] Taung is the type specimen of *A. africanus*, so that binomen automatically adheres to it. To the extent that any other specimens have been incorrectly assigned to its species, any and all reassignments would apply to the others (Kimbel, 1990).

evolutionary interpretations of *A. africanus*, are contentious. Fossils attributed to *A. africanus* have been recovered from four sites, all in South Africa: Taung (Dart, 1925), Sterkfontein (Broom, 1936), Makapansgat (Dart, 1948b), and Gladysvale (Berger et al., 1993). All four are cave remnants, and have been affected by erosion, collapse, and flooding. As a result, the stratigraphy is quite complex, and standard radiometric dating techniques are inapplicable due to the absence of volcanic deposits. This lack, combined with the notorious difficulty of dating speleothems, hinder efforts to achieve consensus dates for the deposits and therefore the fossils. In general, the dates are estimated by comparison of faunal material with that from more securely dated East African sites (e.g. Vrba, 1982), although paleomagnetism is also appropriate because at least some of the relevant sediments formed under water (Rayner et al., 1993). The Sterkfontein depositional bed known as Member 4 is the source of most of the *A. africanus* material, and exemplifies the dating controversy. This member forms the bulk of the area formerly known as its Type Site (Lockwood and Tobias, 2002). It may be as young as 2.5 – 1.5 Ma (Berger et al., 2002), or have a minimum age of over 2 Ma (Clarke and Partridge, 2002; Thackeray and Dupont, 2006).

Further complicating the dating, the electron spin resonance (ESR) dates obtained from a collection of bovid teeth scattered throughout sediments believed to be part of Sterkfontein Member 4 have a bimodal distribution. The peaks of this distribution are at 1.7 Ma and 2.4 Ma (Schwarcz et al., 1994). Those authors note that the deposits had been affected by decalcification, so there may be mixing between Member 4 and the later Member 5, accounting for the recent date of the later peak. They also point out that the geological age distribution for subsamples of individual teeth cannot be explained by

experimental error alone, indicating that variation in radiological dose-rate within the site may be a factor. Deeper in the Sterkfontein deposits, Member 2 contains the "Little Foot" Stw 573 skeleton that has not yet been fully excavated (Clarke, 2008). It is either definitely older than 3 Ma (Clarke and Partridge, 2002), or definitely younger (Berger et al., 2002), and much of this disagreement centers on the interpretation of paleomagnetic data. Similarly, the Makapansgat fossils seem most likely to date to about 3.0 – 2.6 Ma, but also might span the entire period from 4.0 – 2.0 Ma (reviewed in Schwartz and Tattersall, 2005). They are, however, generally accepted as older than most or all of the Sterkfontein specimens (possibly barring Member 2 of Sterkfontein as discussed above). These controversies are far from settled. One group of researchers (van der Merwe et al., 2003), for example, simply note the uncertainty and settle on a middle estimate of 2.5 – 2.0 Ma for the entire *A. africanus* sample.

Phylogeny, the history of discovery, and a taxonomic reappraisal

One of the more prominent considerations of the phylogeny of the fossils attributed to *A. africanus* is the one published by White and colleagues (1981). At that time, their taxonomic unity was generally considered to be settled (following Robinson, 1954), and those authors gave only passing reference to variation among them. They argued that *A. africanus* was ancestral to the entire "robust" clade of hominins (then only *A. robustus* and *A. boisei*), and were met with some degree of agreement (Strait et al., 1997) despite the persistence of contrary arguments (Tobias, 1980, 1989). The announcements of approximately contemporary taxa *A. aethiopicus* (Walker et al., 1986) and *A. garhi* (Asfaw et al., 1999), however, have complicated this interpretation, and

those authors abandoned it shortly after the discovery of *A. aethiopicus* (Kimbel et al., 1988). Even *A. afarensis*, once thought to be the only hominin then extant, appears to have at least one contemporary in *Kenyanthropus platyops* (Leakey et al., 2001), and possibly another in *A. bahrelghazali* (Brunet et al., 1995), although their status as species separate from *A. afarensis* is not without controversy (White, 2003; White et al., 2000). Those discoveries indicate that hominin phylogeny may have included at least two lineages older than *A. africanus*. Either may be ancestral to many or most later species. Therefore, *A. africanus* may not be ancestral to any later taxon. Another of the sources of contention is likely the relative lack of autapomorphies within *A. africanus*: some of the traits taken as definitive include plesiomorphies shared with *A. afarensis*, and derived traits shared with some or all of *A. robustus*, *A. boisei*, and early *Homo* (Lockwood and Tobias, 1999). Many have therefore argued that a reappraisal of *A. africanus* phylogeny is required (Delson, 1987; Johanson, 1989; but see Skelton and McHenry, 1992 for the opposite view).

Perhaps partly as a result of this challenge, some authors have expressed interest in reexamining these fossils' taxonomy. A well-supported alpha taxonomy is necessary for useful phylogenetic analyses (Uchida, 1992), and this phylogenetic difficulty may indicate a taxonomic error, or simply the failure sufficiently to acknowledge a taxon's variability (Tobias, 1989). Clarke (e.g. 1988; 1994) has long suggested splitting *A. africanus* into multiple taxa, and has been joined by authors such as Lockwood (e.g. 1999; 1997), Picq (1990), Moggi-Cecchi (e.g. 1997; 2001), and Plavcan (2003). They argue that variation among these specimens does not match the patterns seen among other hominoids. Similarly, others (Strait et al., 1997; Kimbel and White, 1988; Kimbel et al.,

2004) have found unusual variation within *A. africanus*, and suggest that hominin phylogeny may be clarified if it is shown to comprise two taxa. This possibility is supported by the finding that phylogenies can generally be more easily and accurately constructed among large numbers of taxa (Pollock et al., 2002).

Clarke (1994; 1988) succinctly describes the implications of *A. africanus*'s variation for our understanding of its phylogeny. He points out that the particular specimens and traits on which one chooses to focus necessarily influence the inferred phylogenetic position of the species. Those who focus on the individuals with small, *Homo*-like teeth view *A. africanus* as a *Homo* ancestor (e.g. Olson, 1985). Others place greatest emphasis on the ones that share traits with the robust group of australopithecines and view it as an ancestor of *Paranthropus* (e.g. White et al., 1981). Still others emphasize the range of variation and are therefore likely to see *A. africanus* as a likely ancestor of both *Homo* and the robust australopithecines (e.g. Skelton et al., 1986; Tobias, 1989). Clarke goes on to point out that this high degree of variability may be due to an unrecognized second taxon in the Sterkfontein sample, with the implication that the resulting addition of a new species and reduction of intraspecific variation would simplify efforts to construct a reliable phylogeny. He has contrasted (1988) the large teeth, thin supraorbitals, and flat face of Sts 71 and Stw 252 with the opposite conditions in Sts 5, Sts 17, and Sts 52, and argues that this pattern is quite unlikely to be the result of sexual dimorphism. He remains the most vocal proponent of splitting *A. africanus* into two species with approximately equal numbers of specimens (e.g. Clarke, 2008, 1994, 1988). While his taxonomic hypothesis has not been adopted by most other paleoanthropologists, it should be pointed out that the choice of Sts 5 vs. Sts 71 as

reference has a noticeable effect on reconstructions of Stw 505's overall cranial morphology, while reconstructions of Sts 71 based on either Sts 5 or MLD 37/38 are nearly identical (Gunz, 2005). This result implies that Sts 71 has substantially different morphology from Sts 5 and MLD 37/38, but is not by itself a test of taxonomic hypotheses.

Whether *A. africanus* contains one or multiple species, the specimens' variation is complex. Both the toughness (Scott et al., 2005) and carbon isotope content (van der Merwe et al., 2003) of its diet were more variable than those of any other known hominin, but the isotope variation does not correspond to the variation in tooth size (ibid.). Although some of its linear tooth measurements are bimodal, unlike every other australopithecine species (Kimbel and White, 1988), there does not appear to be strict covariance between tooth size and cranial features (Lockwood, 1997; Wood, 1994; Kimbel and Rak, 1993; Clarke, 1988). This most likely indicates that the observed variation is not exclusively due to sexual dimorphism (Lockwood, 1997; Kimbel and White, 1988; Clarke, 1988), but not all authors agree that the variation is so substantial as to warrant a taxon split (Ahern, 1998; Dean and Wood, 1982; Wood, 1994; McCollum, 2000). Further, even those who agree that the single species hypothesis is probably untenable disagree about which specimens represent a "new" taxon (Lockwood and Tobias, 2002, and others below).

In distinction to Clarke, most authors identify what Lockwood (1997:299) calls "exceptional specimens" that fail to fit well with the Sterkfontein Member 4 and Makapansgat Member 3 samples. In addition to those discussed above, for example, Picq (1990) identifies MLD 37/38 as having robust-like morphology. Kimbel, with White

(1988) and with Rak (1993) say that Sts 19 has unusual morphology when compared to the rest of the sample, but Ahern (1998) strongly disagrees. Kimbel appears to have changed his mind about Sts 19 at some point in the 1980s, as he earlier (in 1984) accepted Tobias's (1967) assignment of Sts 19 to *A. africanus* without comment. He is not alone in this; at least one other researcher has apparently also had second thoughts about it, but in the other direction. Grine included Sts 19 in a *Homo habilis* comparative sample in 2004, but omitted it from any species in a broad hominin sample in 2008 (Strait and Grine, 2004; Smith and Grine, 2008). Even the famous Sts 5 "Mrs. Ples" specimen has a much longer cranial base than the other specimens, which is the primitive condition (Dean and Wood, 1982), although its stratigraphic position indicates that it may be one of the more recent Sterkfontein Member 4 fossils (Thackeray and Dupont, 2006). Others have also examined the possibility of a temporal component to morphological differences among these fossils and argue against it. Morphology of these specimens does not appear to correspond to geological age as evidenced by different sites (with different ages) (Tobias, 1980; Kimbel and White, 1988), or by positions within a given talus cone (Clarke, 1988). Other specimens identified as problematic include Stw 151 (Moggi-Cecchi et al., 1998), Stw 183 (Lockwood and Tobias, 2002; Moggi-Cecchi, 2001; Lockwood, 1997), Stw 252 (Spoor, 1993; cited in Lockwood, 1997), and Stw 255 (Lockwood and Tobias, 2002).

Stw 53 has also received substantial attention. Hughes and Tobias (1977) originally identified the Sterkfontein deposit in which it was found as Member 5 (which contains stone tools that would tend to indicate affinity or at least contemporaneity with *Homo*), but Kuman and Clarke (2000) argue for the existence of a specific "Stw 53 Infill"

deposit, intermediate in age between Members 4 and 5, and probably postdating 2.6 Ma. They note that this deposit does not contain stone tools, but Pickering and colleagues (2000) argue for the presence of cutmarks on the fossil. If the Pickering group's finding is correct, then stone tools were in use at the time the Stw 53 individual lived, and their absence from this particular deposit, or section of the deposit, is merely another example of the sampling problem endemic to paleontology. It would then appear to be the case that Stw 53 was at least contemporary with a stone tool-using population, if not necessarily a member of one. Several researchers (e.g. McCollum, 2000; Lockwood and Tobias, 2002; Curnoe and Tobias, 2006) have joined its discoverers (Hughes and Tobias, 1977) in identifying Stw 53 as *Homo*, but others place it within *A. africanus*, or at least not *Homo* (Braga and Boesch, 1997a; Braga, 1998; Kuman and Clarke, 2000). Further complicating the situation, Clarke (2008) argues that Stw 53 has many similarities to the KNM-ER 1813 group of *H. habilis sensu lato* (Kramer et al., 1995; Miller, 2000), but more importantly that all of these specimens, East and South African, belong to *A. africanus*. Schwartz and Tattersall (2003) note Kuman and Clarke's (2000) dissent, but nonetheless include this specimen with their treatment of *Homo* specimens from Sterkfontein. Stw 53 is viewed here as an early specimen of *Homo* because of its affinity to East African *Homo*, the presence of cutmarks, and the apparent consensus that it postdates all of the Member 4 specimens.

Interestingly, while many researchers support the hypothesis that *A. aethiopicus* is ancestral to all later "robust" hominins (reviewed in Klein, 2009), there are affinities between some *A. africanus* specimens and *Homo*, and between other specimens and *A. robustus* (e.g. Lockwood and Moggi-Cecchi, 1998; Dean and Wood, 1982; Tobias, 1989;

Picq, 1990; Strait et al., 1997; Braga and Boesch, 1997a; Falk and Gage, 1997; Moggi-Cecchi et al., 1998; Sherwood et al., 2002; Kuman and Clarke, 2000; Clarke, 2008). Therefore, Clarke's (2008) argument described above implies that dividing *A. africanus* will modify the perceived phylogenetic relationships among Plio-Pleistocene taxa. Although twenty years have passed since Johanson's (1989) call to clarify the phylogenetic status of *A. africanus*, the issue remains unresolved. Clarification of its taxonomy, which is the goal of this project, should help to resolve the problem of *A. africanus* phylogeny (Strait et al., 1997; Kimbel et al., 2004).

Hypothesis

The hypothesis under examination here is that the sample of fossils traditionally assigned to *A. africanus* contains members of a hidden second taxon. This hypothesis is tested with reference to whether particular specimens, or a group of them, exhibit consistent morphometric and morphological distinctions from the others. The corresponding null hypothesis is monotaxy of these specimens[3], following Robinson's (1954) taxonomy and the arguments of more recent "lumpers" such as Ahern (1998).

[3] Almost all researchers use the single-species hypothesis as the null (e.g. Cameron, 2003; Ahern, 1998; Lockwood, 1997; Cope and Lacy, 1992). Aiello and colleagues (2000), however, have suggested that since some groups of extant species do not differ morphologically while other species are polytypic, it may be appropriate to test dual hypotheses: conspecificity and multiple species. The former approach is adopted here.

Differentiating fossil species

As it is impossible to observe the mating habits and fertility of fossils, there is no direct means to determine which groups of fossils comprise single species according to the Biological Species Concept (Mayr, 1970; Guy et al., 2003). Rather, paleontologists must assess the degree and type of variation a sample of fossils exhibits and decide whether it is consistent with the inference of a single species. The process is complicated by biological reality: fossils represent a very small sample of evolving lineages. Aside from the issue of phyletic gradualism versus punctuated equilibria (Eldredge and Gould, 1972), evolution is a process rather than an event. We should expect intermediate forms (indeed, all life consists of intermediate forms), and nature does not retroactively conform to our labels of convenience. Thackeray (2003) underscores this point with morphometric data indicating that for some specimens at least, the distinction between *Homo* and *Australopithecus* is not as clear as we might hope. If fossil taxonomy held easy answers, paleoanthropologists would only rarely have taxonomic arguments.

No consensus methodology for paleotaxonomy has emerged (see Kimbel and Martin (1993) for a sample of approaches applied to primates). At one extreme of the spectrum of "splitters" versus "lumpers" are scientists like Ian Tattersall, who has argued (in 1986) that because there are separate extant primate species with nearly identical skeletal morphology, any diagnosable distinction between fossils should result in the identification of a new species. Since then, Tattersall has modified his terminology and now identifies "morphs" in fossil assemblages, for later synthesis into species-level taxa that necessarily encompass some degree of variation (e.g. Schwartz and Tattersall, 2005).

At the other extreme are those like Bjarne Westergaard (1989) who argue that the degree of variation within fossil species is unknowable, so species-level distinctions should only be drawn when there is a clear absence of morphological intermediates between two forms. As with the "any diagnosable distinction" criterion, this extreme position seems to defy consistent practical application. There does not seem to be an objective criterion by which one can decide whether a single binary character is sufficient, or a greater number of characteristics or morphometric data would be required.

One approach may be to bring in other lines of evidence, such as the ecological argument inherent in Wolpoff's (1971) hypothesis of a single hominin lineage due to competitive exclusion. Based on expectations derived from the evolutionary species concept that divergent species will have distinct evolutionary tendencies and historical fate, he argued that two related species cannot be sympatric without pursuing different niches, because otherwise they would be locked in a struggle for existence from which only one could survive. One of the tests he proposes for assessing the presence of niche separation is a requirement that the two putative species' distributions of total cheek tooth crown area have no overlap, and that the difference must become more pronounced with time. If one does not accept this criterion, for example arguing that separate lineages are free to evolve separately from one another or might even evolve in parallel (Dean and Wood, 1981), then "the question of separate australopithecine species would become unanswerable, and hence phylogenetically meaningless" (Wolpoff, 1971:609). Wolpoff and his colleagues (1994) no longer make such dire claims even when arguing for a single evolving lineage within *Homo*. They argue instead for an explicit recognition of

polytypic species in the fossil record and an abandonment of the use of chronospecies, favoring species divisions only at cladogenetic events.

The most commonly held intermediate position is that multiple species should be recognized among groups of fossils when their variation exceeds that seen among similar extant species (Uchida, 1992; Albrecht and Miller, 1993; Miller et al., 1997; Miller, 2000; Harvati, 2003; Miller et al., 2004). Such variation can be temporal and/or geographical, and a reference to some variability model, such as an extant species or genus, is required. This approach implicitly adopts the morphological species concept (Benton and Pearson, 2001). It has often been pointed out (Albrecht and Miller, 1997; Trinkaus, 1990) that intragroup variation can fluctuate due to sampling effects, especially with small fossil samples, and that the simple observation of high variability does not logically require rejection of the null hypothesis of a single species— hence the reference to extant variation. This approach is reasonable, although it is not without its own problems (Wood and Lieberman, 2001; Ackermann, 2002). There is no particular reason, especially given the modern situation of diminished ape speciosity (Dolhinow and Fuentes, 1999), to assume that any modern taxon represents the limit of intraspecific hominin variability. Kelley and Plavcan (1998; Kelley, 1993) underscore this point with their argument that specimens previously attributed separately to either *Sivapithecus* or *Ramapithecus* may comprise one species with an unusually high degree of variation, *Lufengpithecus lufengensis*. In addition to varying levels of actual intraspecific variation, taphonomic damage may inflate observed variability and therefore bias taxonomic studies (White, 2003; Silverman et al., 2001).

Tooby and DeVore (1987) point out the practical distinctions between *referential* and *conceptual* models for understanding the past. In their usage, referential models use a real phenomenon that is easily studied, such as one in the present, as a model by which to study another phenomenon for which data are less available, such as one in the past. A conceptual model represents a different approach to making inferences about unobservable phenomena. A model like this is a theoretical framework with defined concepts and variables, among which outside information indicates that the relationships are tightly constrained, and for which all assumptions are explicit and either experimentally validated or at least realistic. They acknowledge that despite their preference for conceptual models over referential ones, it is difficult to construct valid conceptual models.

Those authors make a strong case for the application of conceptual models to efforts to understand past hominin behavior, but it is not clear that valid conceptual models can yet be constructed for analyses such as this one, which test the taxonomic integrity of a putative species in the fossil record. Before a hypothetical second species is identified within *A. africanus,* we can not know its phylogeny: whether it diverged from a population of *A. africanus* as currently understood, or separately from a late population of *A. afarensis*. Further, we can not know the mode of speciation involved: whether it began to pursue a different niche from *A. africanus* in an instance of parapatric speciation or diverged in a different environmental regime (allopatric speciation) and species' ranges alternated over the sites as the climate fluctuated (Strait et al., 2009). These are but two examples of the incompleteness of our knowledge, but the answers to those issues would necessarily affect the parameters of any conceptual model for

morphological variability at these sites. The limitations of a referential model, on the other hand, are more easily understood. If it can be shown that the selection of *Pan* (described below) is likely to be invalid, or that for some other reason a gorilla or orangutan analog would be more appropriate, the model can be modified as soon as relevant data are acquired.

Many researchers consider modern variation to be the best available benchmark against which to compare fossil variation (e.g. Lockwood, 1997; Miller, 2000; Kimbel and White, 1988; Moggi-Cecchi, 2001; Harvati, 2003; Harvati et al., 2004; Guy et al., 2003). Authors using this approach reject the application of an arbitrary limit on variation (Plavcan and Cope, 2001), although a conceptual model that is well-crafted according to Tooby and DeVore's (1987) criteria would be ill-described as arbitrary. The interpretive power of the approach can be further strengthened by using multiple lines of evidence, such as both metric and qualitative data (Plavcan and Cope, 2001), and by taking advantage of the objective comparisons made available under morphometric methods (Guy et al., 2003). In this case, the use of a referential model is the best-supported approach.

Pan *as a model for species-level variation in the fossil record*

As discussed above, the genus *Pan* is used here as the modern variability reference. The three *P. troglodytes* subspecies occupy geographically distinct areas, demarcated by major rivers and/or great geographical distance. *Pan paniscus* is separated from them by the Congo River, and identified as a separate species because of

genetic, morphological, and behavioral distinctions (Fruth et al., 1999; Braga, 1998; Hill, 1969; Coolidge, 1933). Geographic separation necessarily results in some degree of reproductive isolation in the wild, so the inclusion of all three subspecies of common chimpanzees offers a reasonable proxy for the temporal variation contained in the fossil sample (Richmond and Jungers, 1995). At a minimum, it reduces the risk of spurious results produced by a sampling effect likely to result from the use of only one or two subspecies.

Pan was chosen here over other great apes because of its reduced level of sexual dimorphism in the overall cranium relative to gorillas and orangutans (Ahern et al., 2005). The level of sexual dimorphism in *A. africanus* and a hypothetical second species within it probably cannot be determined prior to determining whether that second species existed. There is, however, circumstantial evidence that cranial sexual dimorphism in *A. africanus* is more like that in one of the *Pan* species than that in gorillas or orangutans. Lockwood (1997) has reviewed efforts to diagnose sex for *A. africanus* specimens, and notes extensive disagreement. If sexual dimorphism dominated variation within that group, there should be more agreement. This controversy implies that sexual dimorphism within *A. africanus* was either reduced or at least did not result in bimodal craniometrics, as with gorillas (Ahern et al., 2005)[4].

The level of variation between the two species of *Pan* (*P. paniscus* and *P. troglodytes*) offers an excellent model against which to evaluate variation within *A. africanus* (Harvati, 2003; Miller et al., 2004). The two *Pan* species (reviewed in Braga, 1998) have generally similar morphology, but are distinct with respect to cranial

[4] Although Kimbel and Rak (1993) noted a bimodal distribution for certain *A. africanus* tooth size measurements, they pointed out that this pattern did not correspond to craniofacial variation.

measurements (Shea et al., 1993; Lockwood et al., 2002; Lockwood et al., 2004; Hill, 1969). The same is true of the subspecies within *Pan troglodytes* (*P. t. troglodytes, P. t. verus, P. t. schweinfurthii*) (Shea et al., 1993; Hill, 1969). These are the congeneric pair of species most closely related to hominins.

These three *Pan troglodytes* subspecies, on the other hand, offer variation at a level below species, but which nonetheless corresponds to some degree of reproductive isolation (Fruth et al., 1999). Data from each are included here in order to characterize variation within that species as thoroughly as possible and maintain the multiple-subspecies approach suggested by Uchida (1992). The reproductive isolation between the modern subspecies may serve as a proxy for time depth in the modern reference sample (Richmond and Jungers, 1995). This project thus follows Miller and colleagues' (2004) suggestion of using a reference taxon that: a) is polytypic, with relatively well-understood taxonomy, b) is genetically closely related to hominins, c) is sexually dimorphic, d) is distributed across a wide range, e) has well-studied morphology, f) is known from large samples of sexed documented specimens, and g) has well-known and characterized levels of intraspecific variation. It does not, however, adhere to their suggestion to compare fossil variation to that between localities, demes, etc., as it is unclear how these particular levels of spatial variation should correspond to fossil assemblages' variation across both space and time. This suggestion is especially problematic when one considers the fact that some populations do not interbreed despite having overlapping skeletal morphologies. Wood and Richmond (2000) employed this logic when they concluded that the fossil record is likely to underestimate the number of species (independently evolving lineages) present.

Anatomical region investigated

There exists a tradeoff in the comparison of fossils' morphology. In an ideal world, all comparisons of biological specimens would be based on complete data in order to avoid confounding caused by homoplasy of a particular trait, such as that which possibly occurred with the hominin masticatory complex (Skelton and McHenry, 1992). For purposes of alpha taxonomy of fossils, however, increasing the anatomical coverage in a given study means reducing the number of specimens that can be directly compared, because few are complete. The goal, then, is to use an anatomical region that preserves well, and is likely to exhibit reasonable amounts of interspecific variation with relatively reduced intraspecific variation. For the reasons described below, the basicranium offers an ideal anatomical area for taxonomic assessment. This approach dates to our earliest knowledge of the South African fossil hominins. Dart, in his original (1925) article announcing the discovery of australopithecines, used a basicranial character (the location of the foramen magnum, inferred from the endocast) to make arguments about the Taung individual's locomotion and taxonomy.

This area of the skull is richly endowed with anatomical features (Aiello and Dean, 1990), facilitating the establishment of precise measurement landmarks. The orientation of the petrosal is independent of face shape and flexion of the cranial base, but does have some relationship to growth of the cerebellum relative to other structures (Dean, 1988). It has also been suggested (Lieberman et al., 2000b; Wood and Lieberman, 2001) that the cranial base may be of excellent utility in paleoanthropological analyses: it exhibits relatively little intraspecific variation; it forms in cartilage rather than membrane, implying more direct genetic influence; it reaches adult dimensions before the

rest of the skull; and it appears to have a significant influence on face shape. The use of an anatomical region with reduced intraspecific variation, such as this one, may also help to increase the apparent prominence of interspecific variation, and also may help to reduce confounding due to homoplasy (Seiffert and Kappelman, 2001), because increased intraspecific variability is correlated with increased homoplasy (Wiens, 1995). In contrast, traits with large amounts of intraspecific variation are likely to have reduced usefulness in taxonomy (Lockwood and Tobias, 2002).

To the extent that the basicranium is independent of cranial and dental development, it may exhibit many selectively neutral traits, which makes it an excellent candidate for studies of taxonomy and phylogeny (Trinkaus, 1990). Overall cranial morphology, however, does not appear to correlate well with hominin phylogeny (Collard and Wood, 2000). On the other hand, while adaptive convergences, drift, and other confounding factors may affect the taxonomic utility of certain datasets, Lockwood and colleagues (2004) found that temporal bone morphology carries a strong phylogenetic signal among extant hominids. They point out that the phylogenetic tree indicated by morphometric data derived from the temporal is congruent with genetic results among great apes and humans. The cranial base also appears to be relatively insulated against nongenetic stimuli, at least as compared to features of the face, masticatory apparatus, and postcranial skeleton. This isolation should in turn increase its taxonomic valence (Wood and Lieberman, 2001).

Uchida (1992) recommends the use of features with little sexual dimorphism for taxonomic analyses, and Kimbel and Rak (1993) have found that the glenoid region of the basicranium has little sexual dimorphism in extant hominoids. Kimbel and White

(1988:189) describe the cranial base of primates in general as "notably nondimorphic," but argue that variation in this anatomical region is excessive in *A. africanus*. Veroni and colleagues (2010), however, found weak sexual dimorphism in the foramen magnum and occipital condyles of a juvenile modern human sample, and they review other work (e.g. Holland, 1986; Gapert et al., 2009) indicating that this small portion of the basicranium is dimorphic among adults. This research project involves different landmarks, but at least some of the same anatomical regions considered by Kimbel and Rak (1993), Holland (1986), and by the groups led by Veroni (2010) and by Gapert (2009). It is therefore not clear *a priori* whether this set of landmarks will capture sex-based variation. For this reason, the assumption of reduced basicranial sexual dimorphism is tested here; see the "Sexual Dimorphism" section of Chapter 2 and the "Sex Distinctions in Basicranial Shape" section of Chapter 3.

The strength of the temporal bone in phylogenetic analyses, the reductions in overall variation as well as sexual dimorphism in the glenoid region, the formation of the basicranium from cartilage, and its relative isolation from nongenetic stimuli together argue strongly for the use of the basicranium in paleoanthropological taxonomic studies.

An overview of the analytical approach

For this project, shape variation among all available basicranial specimens of *A. africanus* was related to that between the two *Pan* species. The analysis was conducted in a conservative two-stage process with features intended to address the problem of missing data due to the fragmentary nature of the fossils. The first stage employed 3D

geometric morphometric analysis (i.e. statistical analysis of shape variation). The mere observation of metric variability is insufficient by itself to support an argument for separating fossil taxa (Plavcan and Cope, 2001; Miller, 1991), and even 3D landmark data omit possibly useful information about the nature of the intervening spaces (Richtsmeier et al., 2002). For these reasons, the first stage of the proposed research served to indicate specimens of interest that may differ in taxonomically-important ways from the remainder of the sample. Their patterns of traditional morphological characteristics were then be used to confirm or refute the specimens' distinctiveness. This second stage employed basicranial morphological characters listed as variable in *A. africanus* by Picq (1990), Skelton and McHenry (1992), Kimbel and Rak (1993), and Strait and colleagues (1997) as well as other characters identified for this study. Steps were taken in both stages to minimize the impact of missing data. The approach is described in more detail in Chapter 2.

For the first analysis, the basicranial morphometric variation between *A. africanus* specimens was compared to that between specimens in a sex-, subspecies-, and species-balanced sample of *Pan*. The comparison of fossil variation to that seen between appropriate congeneric species is an appropriate way to evaluate fossil taxonomy (Miller et al., 2004; Harvati et al., 2004; Harvati, 2003; Ackermann, 2002; Moggi-Cecchi, 2001; Ahern, 1998; Lockwood, 1997; Kimbel and White, 1988). Rejection of the null hypothesis here indicates that basicranial morphometric variability among *A. africanus* specimens exceeds a nonarbitrary index of intraspecific fossil variation. This index, however, is not the only one possible. There are many strong arguments in favor of using *Pan* as the modern reference taxon (see above), but no argument has yet been advanced

that precludes the reasonable application of another taxon such as gorillas or orangutans. Further, statistical inferences drawn from measurements do not necessarily produce results with clear biological significance (Albrecht, 1979). For these reasons, results obtained under this phase are accepted as contingent rather than final, and another stage of analysis is used in confirmation.

This second stage involved traditional morphological characters, in an effort to determine whether specimens with unusual *morphometrics* also exhibit an unusual *morphological* pattern. As shown in Chapter 4, potentially disparate specimens within *A. africanus* were successfully identified and then examined for distinctive patterns of nonmetric morphology. Failure to reject the null hypothesis in the first stage would have indicated that *A. africanus*, like *Lufengpithecus lufengensis* (Kelley and Plavcan, 1998), may simply be a single highly variable species. It could then be seen as adapting with time to its ecological situation as a generalist feeder in a variable and gradually drying climate (Strait et al., 2009). This conclusion would have supported Ahern's (1998) argument that intraspecific variation can account for all variation observed among these fossils, and corroborated Robinson's (1954) original taxonomy. The failure to reject that null hypothesis tentatively supports several other authors' conclusions that observed cranial and/or facial variation exceeds that expected for a single species (e.g. Lockwood and Tobias, 2002; Clarke, 1994; Kimbel and White, 1988). Application of the second stage is intended to lend support to the results generated in the first, and demonstrate that the identified specimens exhibit substantial differences from the main body of *A. africanus*. The specific comparisons and techniques involved in both stages are described in greater detail in Chapter 2.

The combination of metric and morphological approaches is intended to produce rejection of the single-species null hypothesis only with strong evidence to the contrary. Tattersall (1986) argues that groups of fossils should be split into multiple species whenever a diagnosable distinction can be identified. It is excessive, however, to require that fossils be nearly identical to be considered conspecific (Kelley and Plavcan, 1998; Ahern, 1998). The construction of a useful hominin phylogeny therefore requires an appreciation of species variability (Kimbel and Rak, 1993; Albrecht, 2000; Miller and Gelvin, 1993). The mere existence of some distinction is insufficient to differentiate fossil species. To accommodate this problem, the two-stage approach here is intended to be somewhat conservative, but still identify groups that may comprise separately evolving populations. If the data indicate that multiple species within *A. africanus* are likely, this conservative approach will enhance the strength of the conclusion.

Chapter 5 compares the disparate specimens identified under both approaches to those suggested by other researchers (e.g. Clarke, 1994; Lockwood and Tobias, 2002). The extent to which these lists correspond will provide a useful starting point for future investigations.

Chapter 2. Materials and Methods

This chapter describes the samples used in this project, as well as the basic methods used for the analyses. The morphometric data used in this project were collected with a portable digitizer[5] (Microscribe G2X model, Immersion Corporation, San Jose, CA) and analyzed with the *PAST* (Hammer et al., 2001), *tps-Small* (Rohlf, 2003b), *JMP* (SAS Institute, 2007), *Morphologika2* (O'Higgins and Jones, 2006), and *Stata* (StataCorp, 2003) software packages. Morphological observations for the second stage of the project were recorded following visual inspection of the fossils by the author.

The samples

The fossil sample

Morphometric landmark data and traditional morphological observations were collected on all available[6] hominin basicranial fossils from the repositories in South Africa: the Department of Anatomy at the University of the Witwatersrand, and the Palaeontology Department of the Transvaal Museum (Northern Flagship Institution). See table 2.1. The main approach for reducing the confounding effects of taphonomic distortion in the fossils was simply to omit the affected morphometric landmarks, especially if they were few in number. This was the case in several fossils. Stw 13, for

[5] As requested by the curators at the University of the Witwatersrand, a nonmetallic (Teflon) stylus was used in the digitizer in order to minimize the risk and severity of damage to the fossils. The same stylus was used to collect all of the morphometric data.

[6] Per request from the curators at the University of the Witwatersrand, the juvenile Taung specimen was not digitized. The Drimolen *A. robustus* material (Keyser, 2000) was under embargo at the time of data collection, so those specimens were likewise not included in this project.

example, preserves landmarks on the sphenoid and on the temporal bone. As a group, however, the set on the sphenoid are translated and rotated as a result of taphonomic warping from their original location relative to the ones on the temporal bone. The larger number of landmarks retaining their original configuration was on the temporal, so that group was digitized and the group on the sphenoid were omitted. Stw 266 is associated with Stw 255 in a group that also includes at least Stw 254, Stw 256, Stw 259, Stw 260, and Stw 263, and possibly Stw 252 (Lockwood and Tobias, 2002), but none of the other fragments preserve more than a few identifiable landmarks, so it is identified here as Stw 266 in order to minimize confusion as to the particular fragment from which data were collected.

Location	**Specimen**	**Species**[7]
University of the Witwatersrand	MLD 37/38	*A. africanus*
	Stw 13	*A. africanus*
	Stw 504/505	*A. africanus*
	Stw 53	Early *Homo*
	Stw 98	Gracile hominin
	Stw 187	Gracile hominin
	Stw 580	Gracile hominin
	MLD 31	Hominin, gen. et sp. indet.
	Stw 266	Hominin, gen. et sp. indet.
	Stw 329	Hominin, gen. et sp. indet.
Transvaal Museum	Sts 5	*A. africanus*
	Sts 25	*A. africanus*
	Sts 26	*A. africanus*
	Sts 71	*A. africanus*
	TM 1511	*A. africanus*
	Sts 19	Gracile hominin
	SK 47	*A. robustus*
	TM 1517	*A. robustus*

Table 2.1. Fossil specimens included in this project, with repository information and typical species designation.

[7] The species designations listed here are simply "typical" assignments, but some follow Lockwood and Tobias (2002).

The *Pan* sample

As discussed in the Introduction, the species and subspecies within modern *Pan* are used here as a model for similar levels of taxonomic variation in the hominin fossil record (Albrecht and Miller, 1993; Harvati, 2003). Data were collected from *Pan troglodytes* and *Pan paniscus* crania from the following locations. The Powell-Cotton Museum in Birchington, Kent, United Kingdom curates collections of *P.t. troglodytes*, the central subspecies of chimpanzees. The Musée Royal de l'Afrique Centrale (MRAC) in Tervuren, Belgium houses specimens of both the eastern subspecies of chimpanzees (*P.t. schweinfurthii*) and bonobos (*Pan paniscus*), including the bonobo type specimen. The Peabody Museum at Harvard University has the western subspecies of chimpanzees (*P.t. verus*). Finally, the Cleveland Museum of Natural History (CMNH) also has an extensive collection of chimpanzee material, but the subspecies assignments of these specimens tend to be less reliable (L. Jellema, pers. comm.) Table 2.2 shows the sample by number of included specimens and location. Not all specimens were sufficiently complete to allow digitization of all morphometric landmarks. Processing damage was evident on many crania, especially on the occipital condyles. Mostly-complete specimens were digitized despite this problem, as the fossils are incomplete as well, and even an incomplete *Pan* specimen can still preserve all of the landmarks visible on a particular fossil. For this reason, the number of *Pan* specimens used in the individual analyses can vary substantially. The sample sizes for particular analyses are reported with the results of specific comparisons in Appendix 1.

Collection	Females	Males	Unknown sex	Totals
Peabody Museum, Harvard	25	35	-	60
Cleveland Museum of Natural History	40	25	-	65
Powell-Cotton Museum	70	22	-	92
Musée Royal de l'Afrique Centrale chimps	23	25	45	93
Musée Royal de l'Afrique Centrale bonobos	26	22	10	58
Totals	**184**	**129**	**55**	**368**

Table 2.2. Overall sample of *Pan* specimens used for this project.

The MRAC curates a number of specimens of unknown sex. While it is certainly possible to diagnose sex with a reasonable degree of accuracy on these dimorphic crania, this diagnosis was not attempted because it was outside the scope of the project.

Balancing the *Pan* samples

This project addresses the question of the taxonomic status of the set of fossils currently assigned to *A. africanus*. If it is the case that there are two species represented in that fossil sample, it is unlikely that the available fossil sample contains equal numbers of specimens by both species and sex. The *A. robustus* sample, for example, appears to contain a marked excess of males (Lockwood et al., 2007). At first glance, therefore, it may seem unreasonable to establish reference *Pan* samples with a particular sex ratio. The sex ratio of the reference sample will have an effect on the resulting variability distribution, and ideally one would use reference samples with demography similar to that of the fossil sample to be tested.

Given the current state of our knowledge, however, the alternatives to balanced samples are less appealing. It would be possible, for example, to construct a great number of samples with different sex ratios among the *Pan* specimens, and then settle on the result provided by the majority of these scenarios. However, it is not clear how one should distinguish appropriate from inappropriate sex ratios. Stated another way, it is not clear that a particular result obtained under even a majority of scenarios is the correct one; the "true" sex ratio among the fossils may more closely resemble one of those in the minority. This problem also affects recent resampling-based assessments of the sexual dimorphism of *A. afarensis* (Gordon et al., 2008).

One may also argue that we do not utterly lack outside information that can be brought to bear here; namely, the sex of some hominin fossils is apparent based on our knowledge of primate sexual dimorphism of size. The large Stw 505 specimen, for example, is most likely to be male (Lockwood and Tobias, 1999), and the diminutive *A. afarensis* specimen "Lucy" (AL 288-1) is almost certainly female (Tague and Lovejoy, 1998). The discussion in Chapter 1, however, shows that sex diagnoses for *A. africanus* specimens are far from settled. Further, as shown in Chapter 3, sexual dimorphism is not strongly expressed in a genus-wide pattern in the basicrania of *Pan* specimens. If a second species is present in this fossil sample, we do not know the sex ratio of the sample, nor the nature of the species' morphological or morphometric differences, nor how the patterns of sexual dimorphism compare between the two. For example, at the gross level, bonobo (*Pan paniscus*) crania exhibit somewhat less sexual dimorphism than do those of common chimpanzees (*Pan troglodytes*), and also have an overall tendency toward paedomorphy (adult retention of juvenile features) as compared to chimpanzees

(Coolidge, 1933; Schultz, 1969). In a hypothetical collection of incomplete fossil specimens with two species following a pattern like this, it may be difficult to distinguish females of the chimpanzee-like species from either sex of the bonobo-like species. This is merely one possible situation, but it illustrates some of the difficulties inherent to an attempt to use the fossils to generate a model sex ratio for the reference samples of *Pan*, with the intention of applying results based on them back to the fossil sample.

One solution would be to test for the presence of consistent basicranial shape differences between male and female specimens in the overall sample of *Pan* crania. If there are such differences, and if they are consistent by species, then it may be possible to extrapolate that result to the fossil specimens, and then to use that information as a starting point for setting the demography of the reference samples. This analysis has been conducted, and the results are presented in Chapter 3. In brief, there is not a consistent pattern of sexual dimorphism in *Pan* crania that can be extrapolated to the fossils.

For the reasons above, the pooled *Pan* samples used for the analyses described below contained equal numbers of specimens by species and by sex, and the subspecies within *Pan* were represented by equal numbers of specimens. When data were available on more specimens of a particular taxon or sex than on another, specimens were randomly excluded until the desired sample constitution was achieved.

Morphometric Analysis

Analysis of the morphometric data proceeded via three-dimensional (3D) generalized Procrustes analysis (GPA)[8] using the *Morphologika2* (O'Higgins and Jones, 2006), *Stata* (StataCorp, 2003), *PAST* (Hammer et al., 2001), and *tps-Small* (Rohlf, 2003b) software packages. The GPA technique offers minimal error among competing methods (e.g. Euclidean distance matrix analysis and Bookstein shape coordinates) with pattern-free bias, especially when there is substantial variability (Rohlf, 2003a). This technique performs computational rescaling, rotation, mirror-imaging, and superimposition of the specimens in such a way that the sum of the squared distances between landmarks is minimized. While unique solutions exist for two-dimensional (2D) applications, the 3D case requires an iterative algorithm as mentioned above (reviewed in Slice, 2005a; Bush et al., 2002; Lockwood et al., 2002; O'Higgins and Jones, 1998).

Morphometric data

The relative three-dimensional locations of 31 ectocranial landmarks described in Table 2.3 and illustrated in Figure 2.1 form the basic data for the first stage of the project. These points were selected with the object of identifying a substantial number of easily defined and replicable landmarks. The temporal bone is favored here due to its strong phylogenetic signal (Lockwood et al., 2004). All landmark coordinates were taken on the best-preserved side, except as discussed below. Some points are adapted from Lockwood and colleagues' (2002) work on the temporal bone, others are standard osteometric points

[8] As this analysis uses landmark configurations scaled to the same size, it employs the "partial" case of generalized Procrustes analysis (Slice, 2005a).

(White and Folkens, 2000), and the remainder were defined specifically for this project. Some of the "new" landmarks are the anteromedial counterparts of landmarks observed at the most posterolateral extent of foramina used by the Lockwood group (Lockwood et al., 2002). While these paired landmarks are non-independent to some degree, there is evidence (Braga and Hublin, 1998) that at least some of them are likely to have independence associated with varying foramen diameters. The cross-sectional area of the carotid canal, for example, correlates very well with endocranial volume in modern humans, and along similar trajectories in early modern humans, Neanderthals, and *Homo erectus sensu lato*. Braga and Hublin were unable to find any such allometric relationship among chimpanzees. Interestingly, for members of the genera *Australopithecus* and *Paranthropus*, including MLD 37/38 and Sts 5, the carotid canal/brain size relationship follows a trajectory different from that of either the *Homo* group or chimpanzees (Braga and Hublin, 1998). Because this trajectory has changed during the evolution of the hominin lineage, and because hominin phylogeny is unsettled, taxonomically- and phylogenetically-important information may be contained in the diameter (and therefore these landmarks) at least for the carotid canal. The presence of this effect has not been established for the other foramina, but two landmarks each are included here for the sake of completeness.

Several techniques were employed to minimize the effects of taphonomic distortion and damage. Data were generally recorded on the left side of the *Pan* specimens except where there was damage; in such cases the least-damaged side was digitized. For the fossils, the best available side was digitized; if both sides were

	Landmark	Reference
1	Hormion	White and Folkens, 2000
2	Basion	White and Folkens, 2000
3	Opisthion	White and Folkens, 2000
4	Apex of the petrous portion of the temporal bone	Lockwood et al., 2002
5	Most medial extent of petrous portion of temporal bone	This project
6	Most anteromedial point on margin of occipital condyle	This project
7	Most posterior point on margin of occipital condyle	This project
8	Most lateral point on margin of occipital condyle	This project
9	Most anteromedial point of foramen ovale	This project
10	Most anteromedial point of foramen spinosum	This project
11	Most inferior point of eustachian process	This project
12	Most anteromedial point of carotid canal	This project
13	Most anteromedial point on jugular foramen	This project
14	Most anteromedial point of styloid pit	This project
15	Most anteromedial point of stylomastoid foramen	This project
16	Most anterolateral point on jugular process of occipital bone	This project
17	Intersection of the infratemporal crest and sphenosquamosal suture	Lockwood et al., 2002
18	Most inferior point on the entoglenoid process	Lockwood et al., 2002
19	Center of articular eminence	Lockwood et al., 2002
20	Deepest point in mandibular fossa	Lockwood et al., 2002
21	Most inferior point on postglenoid process	Lockwood et al., 2002
22	Porion	White and Folkens, 2000
23	Most inferolateral point on tympanic element of temporal bone	Lockwood et al., 2002
24	Center of inferior tip of mastoid process	Lockwood et al., 2002
25	Asterion	White and Folkens, 2000
26	Most posterolateral point of foramen ovale	Lockwood et al., 2002
27	Most posterolateral point of foramen spinosum	Lockwood et al., 2002
28	Most posterolateral point of carotid canal	Lockwood et al., 2002
29	Most posterolateral point of jugular foramen	Lockwood et al., 2002
30	Most posterolateral point of styloid pit	Lockwood et al., 2002
31	Most posterolateral point of stylomastoid foramen	Lockwood et al., 2002

Table 2.3. Cranial landmarks employed in this project.

available and exhibited similar preservation, both were separately recorded (see below for more discussion on this point). A filling-in of landmarks missing from a particular side

with mirror-imaged landmarks from the other side was not attempted, as such an effort would involve the assumption of near perfect symmetry.

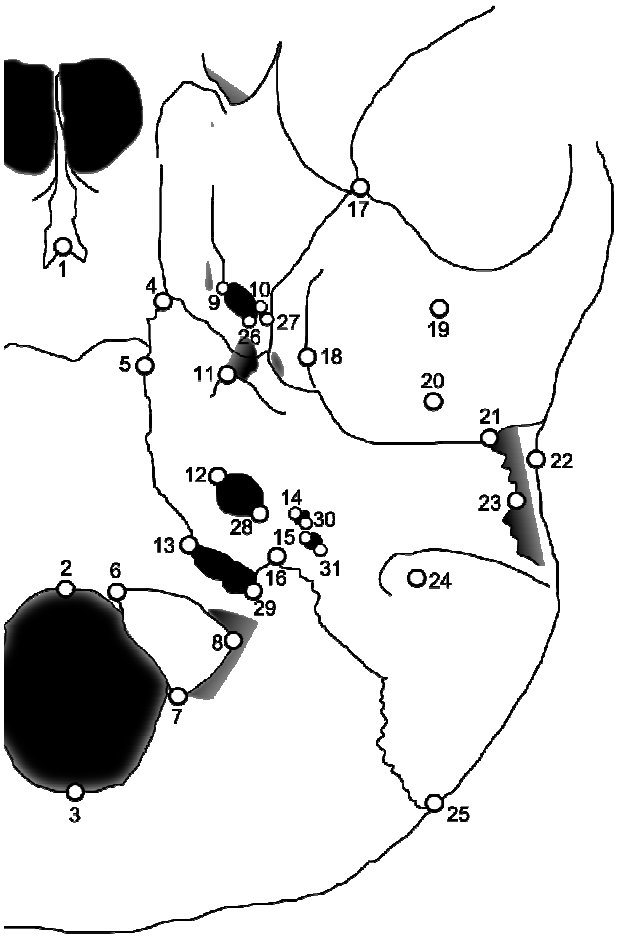

Figure 2.1. Landmarks used in this project (chimpanzee cranium illustrated).

The Procrustes superimposition algorithms used by *PAST* and *Morphologika2* are not sensitive to reflections: these packages automatically (in the case of *PAST*) or by default (in the case of *Morphologika2*) treat the right and left sides of a bilaterally symmetrical object as identical. The *tps-Small* package, however, is sensitive to reflections, so in situations where it was used, the data were first superimposed in another

package and then transferred to *tps-Small*. Procrustes fitting in three dimensions is already an iterative procedure (Slice, 2005a), so simply adding one or a few more iterations does not affect the shape configurations or the following analyses. When centroid size was important, such as for the allometry analysis, it was calculated as part of the first superimposition.

Allometric adjustment to shape

The fossil specimens are not all the same size. Stw 505, for example, is quite large in comparison to the other specimens (Lockwood and Tobias, 1999). Comparing the shapes of biological specimens of different size requires consideration of the nonindependence of size and shape referred to as allometry (Jolicoeur, 1963). The *Pan* sample was examined for the presence of a consistent relationship between size and shape. If such a relationship exists, then some of the shape differences between specimens, especially those of substantially different sizes, is attributable to allometric effects and is therefore not necessarily taxonomically useful. In that case, the shape data for the fossils must be adjusted to remove the size-dependent shape changes and permit comparison of variation that is likely to have more taxonomic importance. The methods by which this step was addressed, and the results of the analyses, are addressed in greater detail in Chapter 3.

Analytical approach, part 1: Procrustes distances

An *a priori* guide to interspecific variation within these taxa was established with reference to the pooled-species sample of *Pan* crania. The distributions and critical values generated for the pooled-species sample offer straightforward indices of species-level variation in hominids, because it contains two related but genetically and morphologically distinct species (Lockwood et al., 1996).

The first method for generating benchmarks for variation involved the distributions of all possible pairwise Procrustes distances (Dryden and Mardia, 1998) between specimens. The data for these histograms were obtained with the *tps-Small* software (Rohlf, 2003b) after an initial superimposition with *PAST* (Hammer et al., 2001) to reflect specimens to the same side as appropriate.

Tests comparing sets of specimens

The Procrustes distances between fossil specimens were compared to the distribution of Procrustes distances between all pairs of a sample of *Pan* specimens in a battery of 21 tests. Three more tests addressed the degree of bilateral symmetry present in three particular specimens, for a total of 24. Table 2.4 shows the groups of specimens included in each test, and Appendix 1 shows details of the landmarks and comparative *Pan* sample included in each. Each set of fossils for which all specimens preserved six or more landmarks in common was tested together, as well as selected groups with four or five common landmarks in order to include as many fossils as possible, with the recognition that comparisons involving relatively few landmarks are less likely to be taxonomically informative. Groups that shared three or fewer common landmarks were

not tested. Each basicranial specimen commonly attributed to *A. africanus* on which four or more landmarks were preserved was included in at least one test, as well as other Sterkfontein and Makapansgat specimens of unknown or controversial affinity, and some exemplar South African specimens from other taxa (either *A. robustus* or early *Homo*). Note that tests 12 and 13 appear to involve the same fossils; they use different sides of the excellently preserved Sts 5 as well as slightly different landmark sets. Also note that three tests (17-19) involve a single specimen each; these were used for the evaluation of lateral asymmetry in the three most complete fossils. This procedure and its results are described in the "Taphonomic Error" section below. While numbered among the others, these three tests are not counted among the 21 in which different fossils' shapes were compared.

For each of these tests, the fossils considered preserve a particular set of landmarks in common. A sample of *Pan* specimens that preserved the same landmarks was then constituted, with balanced species, sex, and chimpanzee subspecies as described above. Details of these samples are presented in Appendix 1. New *Pan* samples were constructed for each of the tests. The *Pan* sample and the fossil sample were then separately subjected to preliminary Procrustes superimpositions in *PAST* to remove reflections, and then all pairwise Procrustes distances between specimens were obtained with *tps-Small*. Each fossil pair's Procrustes distance was then compared to the distribution of pairwise *Pan* distances, and its percentile in that distribution was recorded as a datum for subsequent analyses.

Test	MLD 31	MLD 37/38	Stw 13	Stw 98	Stw 266	Stw 580	Stw 187	Stw 329	Stw 505	Sts 5	Sts 19	Sts 25	Sts 26	Sts 71	TM 1511	Stw 53	SK 47	TM 1517
1		•								•	•					•	•	
2										•	•							
3		•								•								
4		•						•		•								
5		•		•				•		•								
6		•							•		•							
7		•							•	•	•							
8		•								•	•	•		•	•			
9						•	•			•	•							
10										•	•		•					
11		•									•	•						
12		•								•	•	•						
13		•								•	•	•						
14	•	•						•	•	•								
15		•	•							•	•	•						•
16		•			•					•	•							
17		•																
18										•								
19											•							
20		•						•	•	•	•							
21		•						•		•	•							
22		•						•		•	•	•						
23		•							•	•				•				
24		•							•	•	•	•		•	•			•

Table 2.4. Specimens included (•) in each of the morphometric tests.

Weighting tests and calculating interspecimen "distances"

Some well-preserved specimens (e.g. MLD 37/38 and Sts 5) are included in many tests, as they preserve landmarks in common with many or most others. This results in a situation where, for the sake of completeness, a given pair of specimens are compared against one another multiple times. For example, MLD 37/38 was compared with Sts 5 in at least 16 tests involving different sets of landmarks. Test 2 has 24 landmarks representing nearly all of the basicranium, for example, while Test 14 only has four from

the lateral portion of the temporal bone. Tests including greater numbers of landmarks should be more useful than those with fewer. Some pairs of specimens, however, were included in a variety of tests that involved the same number of landmarks, but different sets of landmarks (for example, Sts 71 with MLD 37/38, and Sts 5 with Sts25).

By the same token, some landmarks were not included in certain tests because they were only present in a small number of fossils (such as landmarks 6-8), and some tests included a small number of landmarks in order to address more fossils. In order to take account of the variation expressed through these landmarks and in these fossils, the percentile ranks derived in these smaller tests were also included in the averages. As argued above, however, tests involving fewer landmarks should be given less weight than those covering the basicranium in greater detail. For this reason, when a pair of fossils was included in more than one test, their resulting percentile scores were averaged, with data weighted according to the number of landmarks in each test. Weighted-average Procrustes distance itself was not chosen as the distance metric because distances calculated on different landmark configurations are not directly commensurable. Because of this, the distance proxy chosen for this project was fossil variability, expressed as the proportion of *Pan* pairs separated by a smaller distance in shape-space than that between a given pair of fossils.

This approach, in which the results of several tests of morphometric variability were combined for a given specimen pair, was intended to overcome the tradeoff between including more landmarks and including more specimens in the comparisons, and to overcome at least partially the missing-data problem. Because of the incomplete preservation of the fossils, direct comparisons of fossils can involve either more fossils at

the expense of landmarks, or more landmarks at the expense of fossils. The benefits of this approach are discussed in Chapter 4.

Ordination of interspecimen distances: PCO

With these weighted-percentile data taken as a measure of shape dissimilarity, ordination through principal-coordinates analysis (PCO) with *PAST* was undertaken in an effort to simplify the rather broad collection of inter-specimen distances. Ordination is the general term for a group of statistical approaches to reducing the dimensionality of multivariate data, often for purposes of simplified visualization (Pielou, 1984). Another approach to the depiction of similarities between specimens is cluster analysis. This method was not chosen because the resulting dendrograms can "find" clusters in the data regardless of whether meaningful clusters exist (Jackson et al., 1989), and in general appear to be a reduced-dimensionality form of ordination.

PCO is an eigenvalue-based approach to ordinating a matrix of similarities (or distances) between objects. Like principal components analysis (PCA), it produces a series of uncorrelated components that, together, can be used to ordinate a higher-dimensional dataset. Indeed, applying PCO to a distance matrix typically gives the same ordination as a PCA (Manly, 1994). The advantage of PCO is that it can be applied to similarity matrices that contain some non-Euclidean relationships, as may happen when non-Euclidean distance measures are used.

The PCO algorithm used by *PAST* involves a power transformation with user-selected power c = 1, 2, 4, or 6 applied to the data matrix before centering and eigenanalysis (Hammer et al., 2001). The choice of c value is not trivial, as it can either

mitigate or exacerbate non-Euclidean relationships among the specimens (Podani and Miklós, 2002). In general, higher values of *c* result in higher eigenvalues of the first coordinate, which may simplify visualization, but at the possible cost of the creation of negative eigenvalues indicating departure from the Euclidean condition. When the goal of PCO ordination is the visualization of specimen relationships in either two or three dimensions, a significant departure from Euclidean geometry as evidenced by large negative eigenvalues can seriously confound interpretation (Cailliez and Pagès, 1976; Manly, 1994), but a minor one can be absorbed without major problems (Manly, 1994; Podani and Miklós, 2002). For this reason, the *c* value sought for the ordination was one that yielded both large positive and small (if any) negative eigenvalues, where a "small" negative eigenvalue was defined as one whose absolute value is less than that of the smallest positive eigenvalue (Digby and Kempton, 1987; cited in Podani and Miklós, 2002). The results of this step are presented in Chapter 4 (see section "Principal Coordinates Analysis"). After PCO, the resulting specimen ordination was observed in three dimensions in an effort to identify any specimens that stood in distinction from the rest.

Ordination of interspecimen distances: NMDS

The ordination was also repeated with nonmetric multidimensional scaling (NMDS) (refined by Kruskal, 1964a, b) in *PAST* as a check on the stability of the results seen in PCO. NMDS begins with an arbitrary projection of the original data points in *m*-dimensional space (where *m* is selected by the user; usually 2 or 3), and iteratively nudges the projected points about until the ranks of their Euclidean distances in this

created space correspond as closely as possible to the ranks of their differences in the original dataset. While PCO ordination is typically employed for data with a more linear structure (collections of interspecimen "distances" expressed as percentile ranks are clearly not linear), PCO was selected as the primary data-visualization tool over the nonmetric option of NMDS due to a quirk in the behavior of the latter.

Since NMDS is an iterative optimization procedure, individual runs of the procedure on a given dataset have a tendency to become "trapped" in local minima, where any *slight* modifications to a given configuration result in increased divergence between original and projected distances (a "worse" configuration) and therefore premature termination of the run, but a *substantially* different configuration is better (Kruskal, 1964b). As a result, NMDS is best employed in a series of runs (Minchin, 1987), where its application is simply repeated many times until the researcher is satisfied that the best result obtained is either the best one possible or a close approximation of it.

The foregoing leads one to ask, "but how can one be 'satisfied' of having arrived at the best solution?" One of Kruskal's (1964b) answers was simply to accept a result configuration if it made sense, could be interpreted, or gave the researcher some insight. For a procedure that can provide apparent solutions that may be significantly at odds with the "best" (*i.e.* optimal) one, and for which a unique solution is not guaranteed, this criterion would seem to invite an unsatisfactory level of subjectivity and error. Fortunately, Kruskal offered another in the same publication: the measurement of *stress*, which is an index of disagreement between the set of interspecimen distances as projected in the reduced-dimensional space and the interspecimen differences from the original data (a "badness-of-fit" measure). This index is invariant to apparent differences

in configuration caused by rotation, translation, and reflections as well as scale. This invariance is necessary because the axes in NMDS are completely arbitrary. It is this metric that the algorithm attempts to minimize, subject to the "local minima" traps described above. The *PAST* approach to NMDS selects the lowest-stress ordination resulting from application of the procedure to 11 different starting configurations (one of which is the PCO result) in each run in an effort to evade local minima traps resulting from problematic starting configurations (Hammer et al., 2001), but is still vulnerable to traps resulting from unfortunate iterative "nudges." NMDS runs were repeated 100 times in order to minimize the likelihood of failure to find the lowest-stress ordination. To the extent that they agree, the twin ordination techniques of PCO and NMDS should reliably indicate patterns of shape clustering within the fossil sample.

Analytical approach, part 2: index of variation

The second method of identifying specimens of interest involves Relethford and Blangero's (1990) univariate index of multivariate variability. Although it was originally devised for genetic data, the authors suggested that it could be extrapolated to any multivariate dataset. Their index is derived by taking the trace of the variance/covariance matrix, and dividing this value by the number of traits. This index can be applied to 3D morphometric data when one considers that each landmark is recorded in 3 dimensions, yielding $3k$ variables where k is the number of landmarks. In order to apply their method to these data, size-standardized Procrustes superimposition was first performed on a given group of fossils that shared a particular set of landmarks in *PAST*. A similar

procedure was followed for the overall *Pan* sample, using a new superimposition for each step.

As the variability scores for a given fossil sample are intended to be compared to those of the corresponding *Pan* sample, ratios of the scores were recorded as (fossil variability) / (*Pan* variability). The results were compiled with respect to specimens by taking the natural logarithm of each test's ratio, and averaging the logged ratios for all tests in which a given fossil was included, with weight given to the number of landmarks included in the tests. This approach provides an informal measure of the variability of the overall fossil set when a given fossil is included. Comparing the results for particular fossils should serve as an informal guide to whether any of them introduce excess variation.

This method is neither as sensitive nor as specific as the Procrustes distance method described above, because it is strongly affected by sampling effects. For example, it is possible that a test could include a large number of fossil specimens that are similar and one that is very different. In such a case, the similarity of the majority of specimens may swamp the variability resulting from inclusion of the different one. Alternatively, a specimen with purely typical *A. africanus* morphology may exhibit similar preservation to others with substantially divergent morphology, so it would be tested with them quite often, resulting in variability scores that are high despite the specimen's typical morphology. This variability-index approach is not specific to specimens, so it would not be possible to determine directly whether such sampling effects dominate the tests bearing on a given specimen. Its primary usefulness is in relating the amount of morphological variation in *A. africanus* to that within *Pan*.

Finally, variance can be substantially affected by sampling effects when samples are small, as is often the case here. Therefore, with respect to individual specimens, this technique is used solely as a secondary, informal guide to identify those that may vary from the typical *A. africanus* condition, to confirm that the results of the Procrustes techniques are not entirely an artifact of the analytical approach.

Interpreting the results of the morphometric analyses

The ordination techniques described above were used to identify whether any clear, consistent morphometric distinctions exist among the fossils. Such distinctions would help to indicate fossils with morphology that may be sufficiently disparate to warrant a taxonomic split. Similarly, if the Relethford and Blangero index of variability (1990) for the fossils exceeds the comparable *Pan* index, the fossil sample may be viewed as exceeding the amount of variation in a reasonable modern analog. Conversely, if the fossil sample variation were not to exceed that found within *Pan*, it would be difficult to maintain the claim that the fossils should be assigned to different species. Although it is not a sensitive test, this index can also be used to identify specimens that are often involved in groups of fossils that exceed the variation in *Pan*.

Had no specimens been consistently identified as either having disparate morphology or contributing to an excess of variability within the fossil sample, the null hypothesis of monotaxy would not have been rejected, and the conclusion would have been that the specimens currently assigned to *A. africanus* constitute a single species. As described in Chapter 4, this was not the case.

Since potentially disparate specimens were identified, the analysis moved to the next stage, in which patterns of traditional morphological characteristics were analyzed. The removal of raw size as described above implies that these particular specimens may or may not have been identified by other researchers as illustrative of differences within *A. africanus*.

Morphological Analysis

Variable characters

Upon identification of specimens with unusual basicranial shapes, the next step involved comparison of the groupwise distribution of the morphological characters listed in Table 2.5 and described below. This step is secondary to the morphometric step because of the subjectivity involved in selecting and scoring morphological traits (Falk et al., 1995; Collard and Wood, 2000), but was necessary because morphometric data may not be sensitive to morphological characters (Aiello et al., 2000). The traits included here are only those basicranial characters listed as variable in *A. africanus* by other researchers (including Picq, 1990; Skelton and McHenry, 1992; Kimbel and Rak, 1993; Strait et al., 1997), or identified as varying among the overall fossil sample during data collection. This approach will not bias the results toward rejection of the null hypothesis, because it will only be rejected if the traits covary with one another among the groups identified after the morphological analyses. Adding invariant traits to this analysis would neither modify the level of bias nor affect the outcome (for example, all specimens are expected to have one foramen magnum and two occipital condyles, assuming relevant

	Description	Reference
1	Distance between M^3 and temporomandibular joint	Skelton and McHenry 1992
2	Depth of mandibular fossa	Skelton and McHenry 1992, Picq 1990
3	Position of postglenoid process relative to tympanic plate	Skelton and McHenry 1992, Picq 1990, Kimbel and Rak 1993
4	Orientation of tympanic plate	Skelton and McHenry 1992, Kimbel and Rak 1993
5	Shape of tympanic canal	Skelton and McHenry 1992
6	Angle of petrous bones relative to coronal plane	Skelton and McHenry 1992, Kimbel and Rak 1993
7	Inflection of mastoids beneath cranial base	Skelton and McHenry 1992
8	Vaginal process presence	Strait et al. 1997, Kimbel and Rak 1993
9	Expression of petrous crest	Kimbel and Rak 1993
10	Presence of open mastoid fissure	Kimbel and Rak 1993
11	Relationship of foramen ovale to lateral pterygoid plate	Kimbel and Rak 1993
12	Shape of lateral pterygoid plate	Kimbel and Rak 1993
13	Orientation of medial aspect of tympanic portion of temporal	Kimbel and Rak 1993
14	Size of postglenoid process	Picq 1990
15	Degree of expression of articular eminence	Picq 1990
16	Presence of equilateral triangle formed by postglenoid process, entoglenoid process, and anterior zygomatic tubercle ("tubercule zygomatique antérieur")	Picq 1990[9]
17	Presence of third sulcus on occipital	This project
18	Width of hypoglossal canal gap	This project
19	Rugosity of ectocranial margin of foramen magnum, posteromedial to occipital condyle	This project
20	Angulation of glenoid fossa	This project
21	Presence of blade-shaped process on squamosal part of temporal, near external auditory meatus	This project
22	Presence of "arrowhead" appearance of basilar portion of occipital bone	This project
23	Styloid process or styloid pit	This project
24	Distance between sigmoid sulcus and foramen magnum	This project
25	Presence of eustachian process	Kimbel and Rak 1993
26	Location of sphenotemporal suture relative to entoglenoid process	This project
27	Shape of entoglenoid process	This project

Table 2.5. Basicranial morphological traits.

preservation). The null hypothesis of monotaxy can be rejected if the two groups exhibit generally different morphological patterns, unless there is a corresponding clear

[9] The origin of this trait is in Picq's work, but he only identified the distance between the postglenoid and entoglenoid processes as being equal to the postglenoid process – anterior zygomatic tubercle distance, and only on MLD 37/38.

distinction in site or time horizon (Kimbel and Rak, 1993). If the putative groups do not exhibit different morphology, the null is not rejected and the conclusion is that the fossils assigned to *A. africanus* are likely to constitute a single species. Each of these characters is variable in *A. africanus*. If there are multiple taxa within this set of fossils, it is possible that some of these traits may prove useful in delineating them.

Skelton and McHenry (1992) cite Kimbel and colleagues (1984) as describing the M^3-temporomandibular joint distance as long in *A. afarensis*. This distance is variable in *A. africanus*-attributed specimens, and was recorded as Trait 1.

The mandibular fossa can be either deep, as in modern humans, or intermediate between that depth and the substantially shallower fossa of *A. afarensis* (Trait 2). Both Skelton and McHenry (1992) and Picq (1990) refer to this variation.

For Trait 3, the postglenoid process is either partially merged with the tympanic plate (typically superiorly), or clearly separated from it.

The lateral portion of the tympanic plate can be oriented nearly vertically, such as in *Homo* and the robust group of australopithecines, or intermediate between that orientation and the typically horizontal one of *Pan* (Skelton and McHenry, 1992). This is Trait 4.

Trait 5 refers to the shape of the tympanic canal. Its diameter can expand dramatically from the interior of the temporal bone to the external auditory meatus, or the change in diameter can be much reduced or even absent.

Trait 6 concerns the angle of the petrosals relative to the coronal plane. This angle can be less than 45° as in Sts 19, or it can be 45° or greater as in the bulk of the Sterkfontein specimens.

For Trait 7, the inflection of the mastoid processes below the base of the cranium was scored as strong or reduced.

A vaginal process associated with the styloid process (or pit) can be either present or absent. This variation was scored as Trait 8.

Trait 9 refers to the expression of a petrous crest (Kimbel and Rak, 1993) as either strong or weak.

Trait 10 concerns the condition of the mastoid fissure with respect to being either open or closed laterally.

Variation in the relationship of the foramen ovale to the lateral pterygoid plate is Trait 11; it is either completely posterior to the plate or located adjacent to it so that the plate is indented.

Kimbel and Rak (Kimbel and Rak, 1993) describe a form of variation in the shape of the lateral pterygoid plate. It can be either *triangular*, with its long base superior, or *rectangular*, with its short base superior. This is Trait 12.

The surface of the tympanic twists to a variable degree, such that its lateral and medial portions can be oriented differently. Whereas Trait 4 referred to the lateral portion of the tympanic, Trait 13 refers to the orientation of the medial portion of the tympanic. The medial portion of the tympanic can either face inferiorly or anteroinferiorly.

The postglenoid process is variable in size. For Trait 14, the postglenoid process was scored as either small to moderate (such as MLD 37/38), or large (such as Sts 5).

The articular eminence is also variable. Trait 15 concerns its degree of expression as either slight to moderate (*e.g.* Sts 25) or strong (*e.g.* Sts 71)

The relative locations of the postglenoid process (PGP), entoglenoid process (EGP), and anterior zygomatic tubercle (AZT) can vary. In some specimens, they are arranged as an approximately equilateral triangle. In others, the three distances are clearly unequal. This is Trait 16. As noted in a footnote above, Picq (1990) only identified two distances as being equal (PGP – EGP and PGP – AZT), and only in MLD 37/38. For purposes of this project, the trait as recorded here refers to the presence or absence of rough equality of all three distances. MLD 37/38 does not have the equilateral triangle, but some other specimens do.

Trait 17 refers to the presence or absence of a third sulcus associated with the jugular notch of the occipital, on the neurocranial side of its lateral portion. It is distinguished from the sigmoid sulcus (which would be the "first" groove) and another shallow sulcus connecting the jugular notch to the foramen magnum, running just posterior to the location of the hypoglossal canal. This third sulcus is the most anteromedial of the three where it connects to the jugular notch. The key to identification is that this sulcus travels medially from the anteromedial-most extent of the jugular notch.

Trait 18 refers to the distance from the lateral edge of the occipital condyle articular surface to the nearest edge of the jugular notch, across the opening for the hypoglossal canal. In specimens from Sterkfontein, these edges are almost always parallel. In some specimens, such as Stw 580 and Sts 19, this space is quite narrow (approximately 3-4 mm); on others it is very broad (6-10 mm), *e.g.* Stw 187 and Sts 5. In specimens traditionally attributed to *A. robustus*, the edges of the jugular notch and occipital condyle are not parallel, and the space between them tapers substantially. This

is not simply an age-related trait, as the juvenile specimen Stw 580 and adult Sts 19 have narrow spaces, while subadult Stw 187 and adult Sts 5 have wide ones.

Trait 19 scores the presence or absence of marked rugosity on the ectocranial surface of the occipital, posteromedial to the occipital condyles.

Trait 20 refers to the presence or absence of rotation of the major axis of the glenoid fossa relative to the coronal plane. The glenoid fossa is not a hemispherical pit in any specimens, but rather has a degree of roughly mediolateral elongation corresponding to the shape of the mandibular condyle (Picq, 1990). In specimens exhibiting this rotation, the major axis of this elongation is angled by approximately 15-20° relative to the coronal plane to a more posteromedial – anterolateral orientation. Specimens without it have glenoid fossae that are oriented mediolaterally. This rotation is present in Sts 5, but is not in Sts 19.

Trait 21 refers to the presence or absence of a blade-shaped bony process at the superoposterior margin of the external auditory meatus, oriented anteroposteriorly. It is most clearly seen on Stw 53 as a bladelike spicule 4 mm long and 2 mm high, with a symmetrical low triangle profile. Stw 98 has the same process, though somewhat smaller. It is also present on Sts 5.

Trait 22 refers to the presence or absence of marked depressions on the basilar portion of the occipital for insertion of the longus capitis muscles. The depressions surround the midline pharyngeal tubercle and are bounded laterally by a raised margin. Taken together, the pharyngeal tubercle and the lateral margins form an arrowhead shape. MLD 37/38 offers a strong example. Alternatively, the longus capitis muscles can leave

no depressions, resulting in a much smoother appearance of the basilar portion. The arrowhead shape is absent on TM 1511.

Trait 23 refers to the presence or absence of the styloid process. It is present on MLD 37/38. When the process is absent as in MLD 31, the stylohyoid and stylomandibular ligaments and associated musculature originate from a pit.

In some specimens, the sigmoid sulcus travels very close to the foramen magnum (*e.g.* Stw 580); in others it stays more distant (Stw 187). This variation comprises Trait 24.

Trait 25 is the presence or absence of the eustachian process. This portion of the temporal bone is quite variable. Some specimens have no raised structure at all in this location. Others have only a slight convexity with very low relief. These are the most common expressions in modern humans. Among the fossils examined, some (*e.g.* MLD 37/38) exhibit a truncated ridge, which is a narrow but well-defined ridge or crest covering much of the length of the petrosal and parallel to its long axis. The ridge is barely noticeable at its (postero-)lateral end, but gradually and consistently projects farther from the rest of the bone until it is truncated at its (antero-)medial end. Stated another way, there can be a sharp ridge of bone with an apex in approximately the same location as the finger- or clublike eustachian processes, but its lateral side tapers very gradually toward the surface of the rest of the petrosal. In some specimens (*e.g.* Stw53), this ridge is slightly less elevated and its apex overlaps its "truncated" medial end, creating a tonguelike extension that is also parallel to the long axis of the petrosal. Finally, there is the condition that was scored as *present* for the purposes of this project; a "true" eustachian process that clearly projects inferiorly from the petrosal. Other

expressions (truncated ridge and tonguelike extension) were scored as *absent* because they appear to represent the results of a substantially different regime of soft tissue attachments and resulting stresses. This process is present on Sts 5.

Trait 26 refers to the location of the sphenotemporal suture relative to the entoglenoid process. In some specimens such as Stw 505, it is clearly located near the midline of the process; in others such as Sts 71, the suture is more lateral so that the entire process is located on the sphenoid.

Trait 27 refers to the shape of the entoglenoid process. In some specimens, it is rounded in the sagittal plane and somewhat narrower in the coronal plane to produce a blunt or discoid appearance. In others, the process is narrow in both the sagittal and coronal planes to produce a fingerlike, projecting shape. The degree of independence of this trait from Trait 26 is unclear.

Analytical approach, part 3: morphological observations

All of these traits lend themselves to the recording of data in a binary format. Although there are ordination techniques that allow for mixed data types (such as binary with ordinal), data analysis can be most straightforward when all data are of a single type, especially with substantial amounts of missing data. There are a number of similarity coefficients available for binary data (Cheetham and Hazel, 1969; Kenkel and Booth, 1987; Jackson et al., 1989). The similarity coefficient selected for this project is the "Simple Matching" coefficient, also known as the Sokal and Michener coefficient (Cheetham and Hazel, 1969). It represents the probability that two specimens have the same score on a randomly chosen feature (Kenkel and Booth, 1987), and it is simply the

number of traits on which two specimens have the same expression divided by the total number of observable traits that they have in common, whether similar or not. Because this similarity index is a proportion, it automatically accommodates missing data, as long as specimens share at least one observable trait. The *NTSYS* statistical package (Applied Biostatistics Inc., 1998) was used to calculate these interspecimen similarities.

If data were available on all variables for all specimens, the Euclidean distance metric could be employed. As a measure of dissimilarity, it is robust to a wide variety of applications (Gower and Legendre, 1986). Euclidean distances, however, are dependent on the number of binary variables considered. With $m = 4$ variables, for example, the maximum distance between specimens would be 2. With $m = 9$, the maximum distance would be 3. It is not clear that two specimens with maximum divergence on all observable traits should *automatically* be considered more distinct if they are complete than if they are fragmentary, though the former situation is more reliable. If one takes the position that they should be viewed this way, then the converse situation becomes problematic: the minimum distance is zero, regardless of the number of observations on which it is based. It would be inconsistent to use a distance measure that is sensitive to observation count for different specimens, but not for similar specimens. With the Simple Matching coefficient, distance does not correlate with the number of observations, though one should expect sampling effects when few observable traits are shared between specimens.

After coding the morphological observations, the data (interspecimen similarities) were ordinated with both the PCO and NMDS techniques described above. As described in Chapter 4, however, the PCO ordinations were entirely unsuccessful (all eigenvalues

were approximately equal to zero or simply negative), and the data structure was inconsistent with at least one of the eleven initial configurations in the *PAST* algorithms for NMDS (Hammer et al., 2001), causing software failure. For this reason, the morphological data were ordinated in *NTSYS* with the NMDS approach only. As with the morphometric data, the ordinations were repeated 100 times in order to minimize the likelihood of failure to identify the lowest-stress ordination.

The interspecimen similarities were used raw, without conversion to *Pan*-dataset percentile scores like the morphometric data. Procrustes distances are not directly comparable when they involve different numbers of landmarks, but the Simple Matching similarity coefficient is not so affected by variation in number of observations. Referring the morphometric data to *Pan* variability was required in order to permit comparisons of morphometric data from specimens with different degrees of preservation. This step was not necessary, however, for the similarity scores.

A phenetic approach to taxonomy

This method intentionally disregards characteristics' primitive/derived polarity, because the phylogeny of any possible second taxon within this fossil sample is not presently known. The situation is further complicated by the variability of these character states (indeed, they were chosen for their variability). Others have recognized the difficulty posed by this situation and chosen not to pursue an explicitly phylogenetic strategy for alpha taxonomy, instead preferring a more phenetic approach (e.g. Sherwood et al., 2002; Grine et al., 1996; Cope and Lacy, 1992; Wood, 1991; Kimbel and White, 1988). If polarity were to be considered, the lack of phylogenetic information would

present an unavoidable confound. *A. afarensis* has traditionally been seen as the most likely ancestor of *A. africanus* (e.g. Strait et al., 1997), but with the discovery of the contemporaneous taxon *Kenyanthropus* (Leakey et al., 2001), there are now two known candidates for the ancestry of *A. africanus*, and the possible second species subsumed within it. Although they have not been assigned to *Kenyanthropus*, the teeth from the same and nearby localities at Lomekwi exhibit some similarities to *A. africanus*, such as the frequency of having a protostylid, to the exclusion of *A. afarensis* (Leakey et al., 2001). Even assuming that *A. afarensis* continues to be viewed as the most likely ancestor of *A. africanus*, there are reversals in at least some traits, including the structure of the knee joint (Berger and Tobias, 1996). The early appearance in *Kenyanthropus* of characteristics generally considered as derived indicates that the polarity of many traits used in hominin phylogenetic analyses is now much less clear (Leakey et al., 2001).

Depending on the phylogenetic relationships among these taxa, and their character states, shared traits could be synapomorphies, symplesiomorphies, or homoplasies. Further, variable characters such as these may be more prone to homoplasy (Seiffert and Kappelman, 2001). As the purpose of this project is to clarify the fossils' taxonomy in order that their phylogeny can be better understood, it would be inappropriate to give weight to characteristics based on phylogenetic assumptions that may be shortly rendered obsolete.

Case in point: the enlarged occipital/marginal sinus complex

Substantial enlargement of the occipital/marginal sinus complex of veins draining blood from the brain (O/M) was first noted in East African robust australopithecines by

Tobias (1967). It appears to play a role in thermoregulation of the brain as well as maintaining adequate blood flow during upright posture (Falk and Conroy, 1983; Falk, 1986, 1988, 1990; [but see Braga and Boesch, 1997b, and peer commentaries in Falk, 1990]). This circulatory pattern is present in almost all known members of *A. afarensis* and the robust group of australopithecines, but rare among gracile australopithecines (Falk and Conroy, 1983), although it is present in modern humans (Kimbel, 1984). There is debate about whether O/M is observable in particular specimens such as East African robust specimen Omo L388y-6 (Holloway, 1981; White and Falk, 1999), and it may not be present in the juvenile, fragmentary *A. afarensis* specimen LH 21 (Kimbel, 1984). Conversely, O/M is present in the Taung specimen (Tobias and Falk, 1988). What was once argued to be a situation of absolutes, with traits having gone to fixation in at least the *A. afarensis*/robust group (Falk and Conroy, 1983), now appears to involve different trait frequencies (Falk and Gage, 1998), because both LH 21 and Taung violate expectations.

No student of hominin evolution should be surprised by variation within a population. The O/M example illustrates one of the problems faced by researchers working with small samples of fossils of variable creatures. A particular relationship between trait frequency and hominin lineage could indicate that a given trait carries some taxonomic and/or phylogenetic importance. Nature, however, is an impartial judge of even the most attractive hypotheses. The fact that individuals vary in their expression of many traits, even within a lineage, makes strict phylogenetic strategies for studies of alpha taxonomy problematic when samples are small, as discussed by Sherwood and colleagues (2002). Given the small sample size and the fragmentary nature of many of

the fossils considered here, it is likely that some specimen would preserve an anomalous trait. This is why the present research takes a phenetic approach to taxonomy. Here, a specimen's overall pattern of morphology is considered to be more informative than any particular trait.

Relating results to other work

If the results generated support splitting *A. africanus*, the next step would be to compare the membership of the "new" group proposed here with those listed by other authors, e.g. Clarke (1994), Lockwood and Tobias (2002), Kimbel and Rak (1993), and Picq (1990). As these lists are not themselves identical, there will be a measure of disagreement with at least some of them. The precise nature of the disagreement will provide rich opportunities to better describe the variability of these specimens, whether based in specific, sexual, temporal, or geographic differences. It will also allow a comparison of the relative contributions of facial, dental, and basicranial variation to any overall taxonomic assessment of this sample of fossils.

Sexual Dimorphism

As discussed above, the exact nature of sexual dimorphism in fossil species is often difficult to assess (Häusler and Schmid, 1995; Wood and Quinney, 1996; Tague and Lovejoy, 1998; Reno et al., 2003; Plavcan et al., 2005; Gordon et al., 2008). One group of researchers have even suggested that because there is little reason to assume that

past patterns of sexual dimorphism mimic current ones, it may not be possible to produce reliable sex diagnoses for fossils (Ahern et al., 2005).

The glenoid region of the basicranium shows little sexual dimorphism in extant hominoids (Kimbel and Rak, 1993), and it is sampled heavily in the 3D morphometric portion of data collection. Further, most of the fossils considered here lack the cranial and facial anatomy often used in attempts at sex diagnosis, so the sex of these fossils is not obvious. Veroni et al. (2010), however, were able to discern weak sexual dimorphism in the foramen magnum and occipital condyles in a juvenile sample from Portugal, and review other work (e.g. Holland, 1986; Gapert et al., 2009) indicating the presence of dimorphism in this area among adults. Holland (1986) and the Veroni (2010) and Gapert (2009) groups used different observations than are considered here, so it is not clear that their results should extrapolate to a different set of basicranial data.

Because of the uncertainty of the fossils' sex and the possibility of reliable sex diagnoses raised by this work, it was necessary to attempt to devise a sex-diagnosis technique for the fossils from the *Pan* data. To this end, the morphometrics of the *Pan* sample were evaluated for signs of a reliable diagnostic. Unfortunately, sex plays a small role in basicranial shape among the chimpanzee and bonobo dataset. It is shown in the "Sex Distinctions in Basicranial Shape" section of Chapter 3 that none of the first ten principal components of basicranial shape variation in the pooled *Pan* sample are capable of distinguishing sex. The modern comparison sample therefore does not offer a means to diagnose sex among the fossils, so the results of this project are necessarily contingent on future discoveries of fossil sexual dimorphism of the cranial base.

Measurement error

In any type of quantitative research, it is advisable to measure and report the rates of measurement error: intraobserver, interobserver (if appropriate), and instrument error. This reporting allows an assessment of the degree of precision in the analyses, and therefore offers a gauge of the reliability of the conclusions. One unavoidable and fixed component of measurement error in this study is the Microscribe 3DX's manufacturer-reported accuracy of ±0.23 mm (Immersion Corporation, 1998). This instrument error is, of course, included in all analyses.

The selection of algorithm for measuring and reporting intra- and interobserver error, however, is typically affected by the type of data being collected. When data consist of collections of univariate distance measurements, it could be appropriate to take repeated observations on an individual or group of individuals. One would then report correlations between corresponding measurements, or some other measure of divergence within corresponding sets such as standard deviation, variance, or average deviation from the mean, typically expressed as a percentage of the mean (e.g. White, 1991; Petersen, 2000). In order to maintain maximum comparability with prior work, and because of the type of data, an approach similar to that used by Lockwood and colleagues (2002) was chosen.

This project uses Procrustes superimposition to compare overall shapes instead of analyzing collections of Euclidean distances (linear measurements) between landmarks.[10] Because of this, it is more appropriate to use an error-reporting strategy that deals with landmark configurations (shapes) instead of linear measurements. Lockwood and

[10] It would also have been possible to use Euclidean distance matrix analysis (EDMA) (Lele and Richtsmeier, 1991) to analyze the shapes. Had that been the approach adopted here, an error-reporting strategy of directly comparing collections of univariate measurements would have been appropriate.

colleagues (2002), for example, produced histograms of all possible multivariate distances between pairs of specimens. They then compared the shape variability among iterations of repeated digitizations to that of a sample of available specimens. This approach allows a comparison of the variation caused by measurement error alone to the sum of biological variation and approximately the same amount of measurement error. They demonstrated that measurement error was a small contributor in their case. To this end, the maximum sample of *Pan* specimens for which data were available on all 31 landmarks was used. This approximately sex-balanced sample comprised 98 individuals, as detailed in Table 2.6 below.

Collection	**Females**	**Males**	**Unknown sex**	**Total**
Peabody Museum (*Pan troglodytes verus*)	2	5	-	7
CMNH (*P.t. troglodytes*)	13	11	-	24
Powell-Cotton Museum (*P.t. troglodytes*)	15	10	-	25
MRAC chimpanzees (*P.t. schweinfurthii*)	2	10	7	19
MRAC bonobos (*Pan paniscus*)	10	9	4	23
Total	**42**	**45**	**11**	**98**

Table 2.6. *Pan* sample used for intraobserver error analysis.

Pan dataset intraobserver error

To estimate intraobserver error within the *Pan* sample, one complete specimen from each collection was randomly selected and repeatedly digitized, at least once per day of data collection and never successively. These iterations totaled 15 for each of the

exemplars from the Powell-Cotton, Peabody, and CMNH collections, 10 for the MRAC chimpanzee exemplar, and 11 for the MRAC bonobo exemplar.

The *tps-Small 1.20* software package (Rohlf, 2003b) was used to calculate the Procrustes distances between all possible pairs of specimens in the *Pan* sample, and the same distances between all possible pairs of iterations for each of the five exemplar individuals separately (i.e. not between exemplars). A histogram of these distances for visual comparison was then generated with The *PAST* package (Hammer et al., 2001). There are 4753 pairs of *Pan* specimens, and 415 pairs of exemplar iterations.[11] The resulting figure (Figure 2.2) is presented below.

The maximum Procrustes distance between iterations of digitization for the five exemplars is 0.0833, and the minimum distance between pairs of *Pan* specimens is 0.0910. The mean distance between *Pan* specimens is 0.1418. There being no overlap between these distributions, it is reasonable to interpret reported Procrustes distances between specimens as indicative of actual shape differences as opposed to measurement error. The maximum possible Procrustes distance is $\pi/2$, or 1.5708, so the maximum inter-iteration distance here of 0.0833 is quite small.

[11] The number of possible pairs is the number of combinations (without replacement) of p specimens from a sample of size n, or $C_p^n = \frac{n!}{p!(n-p)!}$. In this case, since $p = 2$, the right side of the equation simplifies to $\frac{n(n-1)}{2}$. Here, for the entire *Pan* dataset, $n = 98$. For the exemplar specimens, n = the number of digitizing iterations (10, 11, or 15).

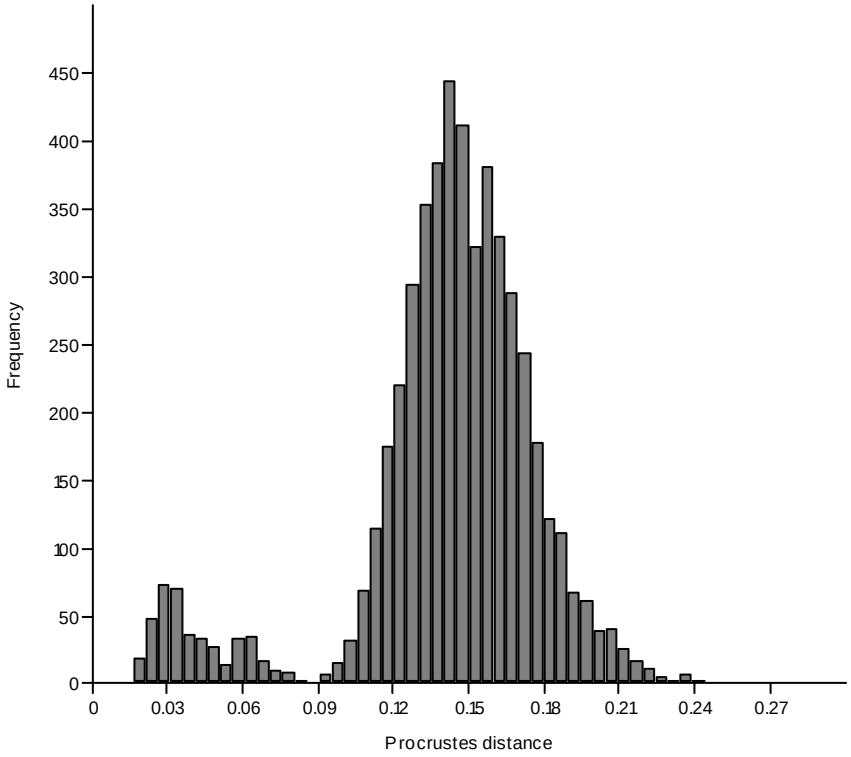

Figure 2.2. Procrustes distances between all possible pairs of *Pan* specimens, and between all possible paired iterations of five exemplar specimens (repeated iterations at left; overall sample distances at right). See text for discussion.

Fossil dataset intraobserver error

As most of the fossils observed here were fragmentary, there were few fossil specimens with sufficient available landmarks and without obvious distortion to make informative descriptions of intraobserver error. The left side of the MLD 37/38 specimen, however, preserves 28 of the 31 landmarks (landmarks 6-8 being absent). This specimen was digitized 10 times. As with the *Pan* exemplars, iterations were recorded on each day of data collection at the University of the Witwatersrand, and never successively. The same *Pan* dataset as used for the *Pan* error analysis for this set of 28 landmarks was reconstituted and the analysis proceeded as above. It was necessary to

present this part of the analysis separately because the shapes being compared are not the same when landmarks are omitted. Again, there are 4753 pairs of *Pan* specimens, and 45 pairs of iterations on MLD 37/38. The resulting histogram is presented below as Figure 2.3.

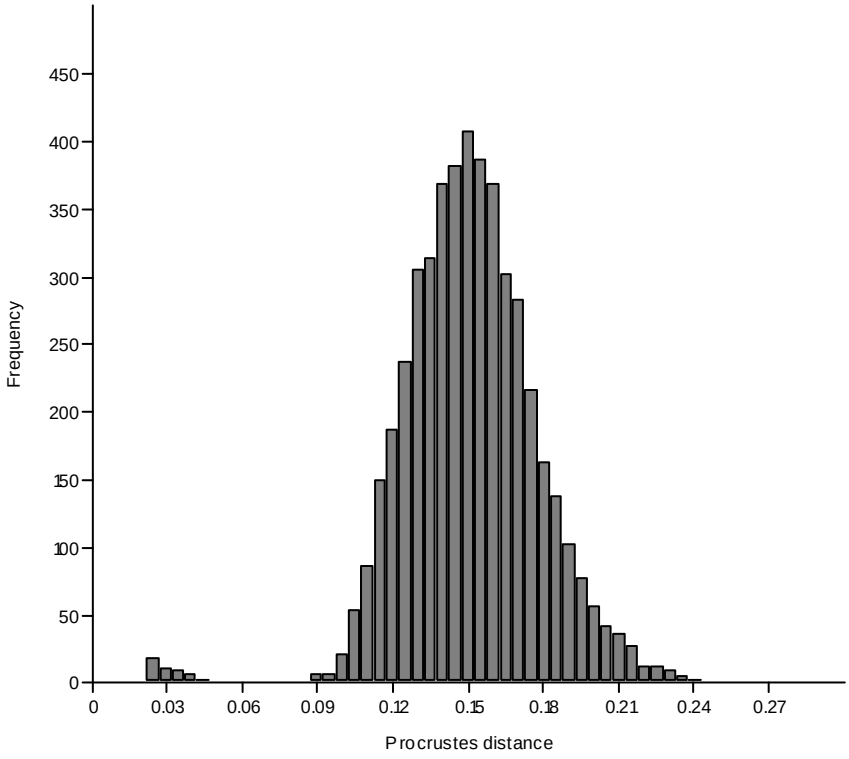

Figure 2.3. Procrustes distances between all possible pairs of *Pan* specimens, and between all possible paired iterations of MLD 37/38 (repeated iterations on MLD 37/38 at left; pairwise *Pan* distances at right). See text for discussion.

Here, the maximum Procrustes distance between digitization iterations on MLD 37/38 is 0.0421. The minimum distance between pairs of *Pan* specimens is 0.0887, and the mean is 0.1521. The distinction between repeated measures and actual specimen differences is even larger here than in the *Pan* dataset. It appears that intraobserver error, at least on MLD 37/38, contributes only a minor amount to observed variation.

It must be noted that intraobserver iterations were not attempted on other fossil specimens. The considerations behind this decision were partly a desire to minimize handling of fossils, and partly to include as many landmarks in the error analysis as possible. This result can be generalized to the other fossils insofar as MLD 37/38 is representative of the clarity (as opposed to extent) of preservation of the other fossils. As stated above, if a landmark was not judged to be clearly visible on any fossil specimen, it was not digitized. The data available reflect easily identified landmarks, so their visibility and clarity should not constitute a major factor in measurement error.

Measurement error conclusion

In no case does the Procrustes distance between repeated digitizations of the same specimen exceed the distance between any pair of *Pan* specimens. Stated another way, measurement error (comprising both instrument error and intraobserver error) does not cause repeated observations of the same specimen to falsely appear to be more distinct than any two individual specimens. It is therefore reasonable to infer that measurement error is not likely to be a major confounding factor in the analyses to follow.

Taphonomic error

In addition to instrument and observer error, the geological processes to which the fossils have been exposed create some error through taphonomic warping. Some of it was alluded to above in the discussion of landmark selection. One way to test the

importance of taphonomic warping as a factor in apparent fossil shapes, at least for a selection of specimens, is to characterize the difference in shape between right and left sides of the same fossil specimen, and to compare that difference to the distribution of shape differences in a reference population. While no specimen is likely to be perfectly symmetrical in life, any taphonomic warping would only reduce the amount of symmetry. Only in the very unlikely case of exactly symmetrical distortion would taphonomic warping be missed with this approach.

The listing below for each test includes the specimen considered, the landmarks present, and the membership of the *Pan* comparative samples. Also listed are the values of Relethford and Blangero's (1990) univariate index of multivariate variability. In this case, the fossil "sample" is the two sides of a given fossil, with one mirror-imaged for superimposition. Finally, the Procrustes distances between specimens are shown, together with the percentiles of these distances in the overall *Pan* and chimpanzee-only distributions of pairwise Procrustes distances. This approach does not take *in vivo* asymmetry into account, but rather serves to highlight cases in which a given fossil exhibits such a degree of asymmetry that the possibility of taphonomic damage cannot be ignored.

Specimen: MLD 37/38

Landmarks present on midline and both sides: 1-4, 9-16, 18-23, 26-31

Pan sample: Chimps: 14 Western, 14 Central, 14 Eastern; Bonobos: 42

Relethford and Blangero (1990) variability index: all *Pan*: .000161, fossil: .000081

Procrustes distance between sides and its percentile in all-*Pan* distribution: 0.1077, 3

The Procrustes distance between the right and left sides of MLD 37/38 is smaller than the corresponding distance between all but about 3% of *Pan* pairs, and the variability index likewise is about half of that for the *Pan* sample. While all of the shape comparisons involving MLD 37/38 used its left side, it is reassuring to note that the difference between the two sides is quite small, consistent with negligible taphonomic warping of its landmark configuration.

Specimen: Sts 5

Landmarks present on midline and both sides: 1-4, 6-10, 12, 13, 15, 18, 20-23, 26-29, 31
Pan sample: Chimps: 12 Western, 14 Central, 14 Eastern; Bonobos: 40
Relethford and Blangero (1990) variability index: all *Pan*: .000144, fossil: .000098
Procrustes distance between sides and its percentile in all-*Pan* distribution: 0.1136, 16.5

In the case of Sts 5, the Procrustes distance between its right and left sides is slightly increased relative to MLD 37/38, such that this difference exceeds those of about one in six pairs in the *Pan* sample. The Relethford/Blangero index for the fossil, however, is well below that for the overall *Pan* sample, as was the case for MLD 37/38. The fossil's left side was used for almost all of the shape comparisons, with a few exceptions as noted in Appendix 1. Because of the small but non-negligible difference between sides, any distinction between the results generated by the two sides must be viewed critically. The most direct examination of the importance of this difference is afforded by Tests 12 and 13 (see Appendix 1), in which the two sides of this fossil were independently compared against the same group of specimens: MLD 37/38, Sts 19, and Sts 25. The landmarks employed in those tests differ slightly. Landmarks 4, 9, 10, 12,

18, 20-22, and 26-28 were common to both tests. Test 12 also included landmark 19, while Test 13 included landmark 11. These varying landmark sets resulted in almost no change in any of the observed distances involving MLD 37/38, but a slightly increased distance between Sts 19 and Sts 25 (0.1649 vs. 0.1540) when landmark 19 was included and #11 excluded, as in Test 12. The specific landmark configuration of Test 12 also increased the Sts 5 – Sts 19 and Sts 5 – Sts 25 distances. Taken as a whole with the fact that the analyses were dominated by Sts 5's left side, these results indicate that the difference between the two sides was unlikely to have contributed substantial error to the project.

Specimen: Sts 19

Landmarks present on midline and both sides: 1-4, 6-16, 18, 26-31

Pan sample: Chimps: 14 Western, 14 Central, 12 Eastern; Bonobos: 40

Relethford and Blangero (1990) variability index: all *Pan*: .000178, fossil: .000092

Procrustes distance between sides and its percentile in all-*Pan* distribution: 0.1101, 3.5

As with MLD 37/38, the Procrustes distance between the two sides of Sts 19 is relatively small, and all comparisons were based on the left side. Here too, the symmetry of the two sides is consistent with a relative lack of taphonomic warping on this specimen.

Taphonomic error conclusion

For both MLD 37/38 and Sts 19, the presence of minimal asymmetry indicates that for these two fossils at least, taphonomic warping appears to be a minor issue. The increased difference between the two sides of Sts 5 may be problematic, but Sts 5 did not appear to be a particularly unusual specimen, and the relationships of other fossils did not appear to be overly dependent on the side of Sts 5 included in the analysis (see above).

Chapter 3. Allometry, Sexual Dimorphism, and Species Differences in the *Pan* Sample

The overall goal of this dissertation is to characterize the extent of shape variation among basicranial specimens assigned to *A. africanus*, in an effort to shed light on the problem of their taxonomy. Regardless of whether there are one or multiple species within *A. africanus*, some proportion of the shape variation between fossils may be attributable to the common biological relationship between size and shape that is known as allometry. It is often the case that larger members of a species have a different shape than smaller ones because not all anatomical structures grow at the same rate, or for the same period of time (Shea, 1983a). This chapter addresses the issue of whether it is appropriate to perform the shape comparisons considered here with a straightforward Generalized Procrustes Analysis (GPA) technique (Slice, 2005a; Dryden and Mardia, 1998). This approach would compare specimens' shapes directly, without taking size-based shape differences into account. The alternative would be that a size-dependent "adjustment" should be made to the shapes of the fossil specimens in order to account for allometry and permit fossil comparison that is not confounded by the shape changes expected to be brought on by differences in size (e.g. McNulty, 2004). An approach such as McNulty's may be thought of as analogous to analyzing the residuals after a linear regression technique has been applied, in an effort to work with variation not "explained" by the linear relationship. Here, the goal of "removing" size-dependent shape variation would be to increase the proportion of variation attributable to species differences.

Unfortunately, the available *A. africanus* specimens are few, fragmentary, and sometimes taphonomically distorted, making it difficult at best to detect and characterize an allometric trajectory through shape-space for them (Cobb and O'Higgins, 2004). Further, if this taxon does comprise multiple species, the possibility of multiple trajectories would render the characterization even more difficult and likely impossible with the currently available sample. Given that limitation, it becomes necessary to examine shape variation in another taxon for which more specimens are available. The two species within *Pan* are generally allometrically scaled variants of one another (Corruccini and McHenry, 1979; Giles, 1956; McHenry and Corruccini, 1981), but this pattern does not extend to all anatomical regions, nor to overall body proportions (Shea, 1983b). The extent to which the basicranium is or is not included in this general pattern is not yet clear, so this chapter includes a test of the extent to which allometry is expressed in adult *Pan* basicrania.

If a strong and consistent pattern of static allometry for chimpanzees and bonobos can be identified in the relatively restricted anatomical region of the cranial base, then it is likely to be the case that allometry was also present in the cranial bases of *A. africanus*. Any efforts to characterize the variation in *A. africanus* cranial bases, then, should take an allometric trajectory through shape-space into account, and that of the two *Pan* species should offer a reasonable estimate. If adult members of the two species of *Pan* do not display an allometric signal in the basicranium that is both strong and consistent, then it could be argued that: a) basicranial allometry related to sexual dimorphism is likely to have been reduced in *A. africanus* as well, and/or b) without a particular allometric trajectory to take into consideration, it would be inappropriate to force an arbitrary one

onto the pattern of variation in the fossils. In the latter case, the results of any shape comparisons should be examined carefully for any size effects. This basic approach has also been used by Bush and colleagues (2002).

Materials and methods

Sample construction and landmark selection

In order to characterize basicranial shape variation in adult *Pan,* sex-balanced adult samples of 38 Eastern (*Pan troglodytes schweinfurthii*), 44 Central (*P. t. troglodytes*), and 12 Western (*P. t. verus*) chimpanzees, and 40 bonobos (*Pan paniscus*) were selected. These are the maximal sex-balanced samples with reliable taxonomic identifications for complete specimens in the *Pan* dataset. Basicranial shape was characterized with 28 of the original 31 three-dimensional landmarks shown in Chapter 2. Three landmarks were omitted for the reasons outlined below.

Landmark 5, the most medial extent of the petrosal, was among those omitted. Since chimpanzee petrosals are relatively coronally oriented, variation in the shape of the posteromedial border of the petrosal resulted in a situation where the "most medial extent" was not homologous between specimens. This problem rendered the landmark seriously deficient for purposes of 3D morphometrics, as it embodies all of the problems with Type 3 landmarks (Bookstein, 1991). Landmark 5 was morphologically problematic because it was sometimes at the tip of the petrosal, near the spheno-occipital synchondrosis, and in other specimens was 1 cm or more posterior to that location, at a

point of curvature along the medial edge of the petrosal.[12] As Bookstein (1991) describes, such a landmark is mathematically deficient as well. A type 3 landmark represents a local maximum in a particular direction, and therefore one dimension is a mathematical function of the others. Differences between specimens at type 3 landmarks have substantial ambiguity in the dimension(s) perpendicular to the "ruler," and thus are reliably indicative only of variation in the length of the defining segment, *i.e.*, a single dimension (Gunz et al., 2005). Other types of landmarks contain useful information in three dimensions. This situation is clearly inappropriate for 3D morphometric analysis of biological shapes, so the offending landmark was omitted.

Landmarks 17 and 25 were omitted because they were not visible on many of the specimens. They are defined partially or entirely by sutures, but as the specimens considered here were adults, the sutures were often obliterated due to age-related changes. Omitting these landmarks greatly increased the available sample size.

Procrustes superimposition

For the analysis itself, the PAST software package (Hammer et al., 2001) was used to perform Procrustes (Generalized Procrustes Analysis; GPA) superimposition of the pooled *Pan* sample, followed by subtraction of the mean landmark configuration, to obtain Procrustes residuals (Lockwood et al., 2002), described below, for all landmarks on all specimens. See Slice (2005a) and Dryden and Mardia (1998) for discussion of the

[12] Cases in which particular landmarks are highly variable between specimens or populations can theoretically be handled with "resistant fitting," which is a median-based approach to shape analysis, where GPA is least-squares based. This approach, however, is not as mathematically robust as Procrustes analysis, and Procrustes methods are preferable (Slice, 2005a). In any event, the non-homologous nature of the recorded locations is fatal.

history and mathematics of Procrustes analysis. Many other works summarize the application of Procrustes approaches to anthropological questions (e.g. Slice, 2005b; Lockwood et al., 2002). It should be noted, however, that the Procrustes residuals are not necessarily devoid of size information, as shape and size may not be independent for these landmarks (Lockwood et al., 2002; Singleton, 2005). To the extent that size and shape are related, the "removal" of size in a single step cannot be complete, as residual size information would still be present in the shapes. The degree of size-shape independence in *Pan* basicrania is the focus of this chapter.

Procrustes residuals differ from raw 3D configurations (the original coordinates in which size, shape, orientation, and overall specimen location information are maintained) and Procrustes configurations (in which shape information is maintained for all specimens, but they are scaled to unit centroid size, rotated to the same orientation, and translated to be superimposed) in that the landmarks no longer maintain their locations relative to one another. After this transformation, the landmarks are recorded as individual perturbation vectors from the mean configuration (the average shape for the sample). This procedure is analogous to one of subtracting the mean from a sample of univariate data, which transforms observations so that they are expressed relative to the mean.

Size proxy

In multivariate studies of allometry, there is no hard-and-fast rule for the way in which size should be controlled for, and the choice of size-standardization variable can affect the resulting allometric equations (Smith, 1981). For a brief review of these issues,

see Singleton (2005). Body mass is perhaps the most common (e.g. Corruccini, 1983; Gauld, 1996), but not the only, size proxy. The selection depends largely on data availability and the anatomical region under investigation. For studies of limb robusticity, for example, it may be appropriate to scale bone diameters against lengths, whether directly or through a logarithmic relationship (e.g. Carlson et al., 2007). It is also possible to use the geometric mean of all variables (e.g. Jungers et al., 1995), but this approach is probably not useful in Procrustes approaches. Often, allometric studies of cranial shape use a measure of neurocranial length, such as glabella-opisthocranion length (Wood and Stack, 1980), but see below for discussion of confounding that can occur in this situation.

For 3D shapes, the use of *centroid size* offers the opportunity to compare shape with a univariate measure of size that directly takes the entire region into account. Centroid size is the square root of the sum of all squared Euclidean distances from landmarks to the centroid, or center of mass (Niewoehner, 2005). The use of centroid size increases the precision of size comparisons, and is uncorrelated with any particular linear distance between included landmarks (Bookstein, 1991) or shape when landmark error is random (Slice, 2005a), making it an ideal overall size proxy for a given set of landmarks (Singleton, 2005, 2002). In particular, its failure to correlate with any linear distance renders it immune to the confounding effects that will accrue when a linear size measurement includes an anatomical feature whose size varies allometrically with body size. This confound could obscure actual relationships (Gauld, 1996).

Body mass was rejected as a size measure for several related reasons. It might be possible to estimate body size from observations made on these fossils (via regression),

but there would have to be many such estimators because of the nonoverlapping preservation of many of the fossils, and they are unlikely to have similar error magnitude or bias. Further, if there are multiple species in this fossil sample and the extent of encephalization is different between them, a body mass based approach to specimen size would be confounded by the absence of prior information about fossils' group membership. Because postcranial adaptations such as limb size and proportions, muscularity, gut size, and adiposity are likely to be independent of basicranial shape, body mass may or may not correlate well within species with overall bone growth in the basicranium.

For these reasons, each specimen's centroid size was recorded (as part of the initial Procrustes superimposition and before the conversion to Procrustes residuals) for use as the univariate size proxy. This approach ensures that any size-shape relationship in the basicranium is kept as direct as possible.

It is common in statistical approaches to biological problems to log-transform the major size variable against which other traits are compared (*e.g.* Singleton, 2005; McNulty, 2004; but see Corruccini, 1987). This is done primarily because the distribution of biological measurements is often not normal (generally having a long "right tail," as it is more often the case that some specimens are unusually large than unusually small), and thus violates a major assumption of most frequency-based statistical approaches. Without this transformation, the larger specimens become outliers and bias the results, especially in regression analyses (Neter et al., 1996). The *Stata 8.2* (StataCorp, 2003) software package was used to perform the Shapiro-Wilk test of normality on the distributions of centroid sizes of chimpanzees and bonobos. In both

tests, the null hypothesis of normality was not rejected[13], so for the purpose of comparing shape against size, analysis proceeded with raw, rather than log-transformed, centroid sizes.

Variable reduction: Principal Components Analysis

Given that 28 landmarks in 3 dimensions yield 84 separate variables (but only 77 degrees of freedom) (Slice, 2005a), it would be convenient to be able to deal with fewer variables while still maintaining access to the shape variation expressed in this dataset. Principal Components Analysis (PCA) (Manly, 1994) is an ordination procedure that reduces a set of original variables to uncorrelated indices known as *principal components*, consisting of combinations of the original variables (*eigenvectors*), scaled such that the sum of the squared coefficients of an eigenvector is 1. With this approach, one can generalize allometric relationships among more than two variables (Jolicoeur, 1963). In cases such as this one, where there are many possible bivariate combinations of variables, PCA is especially well suited for providing a summary (Corruccini, 1983).

A component's *eigenvalue* represents the variance of the dataset as "seen" through that component's particular combination of the original variables. Components with higher eigenvalues capture larger portions of the dataset's overall variance. In this way, the first few principal components may suitably describe most or all of the variance in the dataset, and reduction in variables can follow without serious loss of information. When the original variables are highly correlated, a small number of PCA eigenvectors

[13] For chimpanzees, the raw *p* value on this test was 0.25; for bonobos, 0.44. We cannot reject the null hypothesis of normality, with or without a Bonferroni or Šidák multiple-comparisons adjustment (Wright, 1992).

can capture most of the variance in the original dataset. If they are entirely uncorrelated, the eigenvalues for all eigenvectors will be similar, so PCA will not be able to achieve any reduction in variable number (Manly, 1994), or any distinctions in the relative "importance" of the principal components.

When dealing with raw biological data (whether morphometric or a collection of traditional length measurements), the individual principal components can be interpreted as characterizing particular aspects of shape variation. The pattern of positive and negative coefficients on the eigenvector is unique for each principal component. Each of these essentially represents a comparison of the original variables with positive coefficients against those with negative coefficients. Specimens with large values in variables with positive coefficients and small ones for those with negative coefficients will have large scores for a given PC, and vice versa.

Typically, though, when analyzing raw biological data one of the earlier components (usually the first) will have coefficients that are all positive, while the others involve a mix of positive and negative coefficients. A similar situation often occurs in factor analysis and canonical correlation analysis (Taylor and DiBennardo, 1980). This component is generally taken as representing the overall size of the specimens. As discussed above, however, Procrustes analyses involve an effort at removing size information, and further, the present analysis involves Procrustes residuals. A specimen with consistently "large" Procrustes residuals would not necessarily have a large size, but rather indicate a data problem. In this case, an eigenvector consisting entirely of positive coefficients (or a pattern in which the *x*, *y*, and *z* coordinates for each landmark had consistently negative or positive coefficients) would indicate some serious problems,

namely that the specimens had been arranged linearly through space rather than superimposed, and/or that the scaling to "remove" size information was not successful. As expected, this was not the case, indicating that raw size and faulty superimpositions were not a factor in the observed variation.

As discussed above, this PCA run is based on Procrustes residuals for 28 landmarks on a sex-balanced sample of chimpanzees and bonobos, for a total of 84 raw variables (28 landmarks in 3 dimensions). The eigenvalues for the principal components (PCs) and the percentage of the overall variance that they "explain" are listed in Appendix 2.

The Scree plot offers this information graphically, by plotting eigenvalues against principal component numbers (Figure 3.1). When application of PCA has been successful (i.e., there was a degree of correlation among the original variables underlying the PCA analysis), the plot takes on an L shape, indicating that the first few principal components capture the bulk of the original variance. In such cases, it is reasonable to proceed with a reduced number of variables. "Unsuccessful"[14] PCA runs, as discussed above, exhibit a nearly horizontal line with a more consistent slope. As discussed above, this would happen if the original variables were nearly or entirely uncorrelated, so combining them into principal components could not concentrate the variance into only a few.

[14] The notion of "success" in PCA analyses, or any other statistical procedure, depends on the researcher's goals. If the goal is simply description of data, no situation short of software crash and data loss constitutes failure.

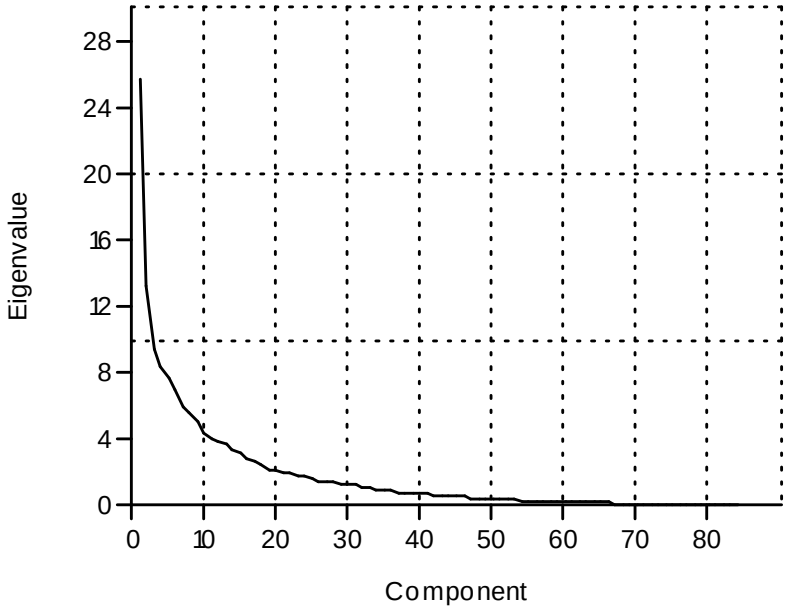

Figure 3.1. Scree plot of eigenvalues against principal components.

It can be seen from Appendix 2 and Figure 3.1 that the first ten principal components account for 60% of the variance in the Procrustes residuals, and that the higher-numbered components have low eigenvalues. The Scree plot is decidedly L-shaped, indicating a successful effort at variable reduction. Given the success of the PCA analysis, it is therefore reasonable to select the first ten PCs as a convenient way to describe shape variation in *Pan* basicrania. The eleventh PC, for example, accounts for about 2.5% of the overall variance, and subsequent PCs cover even less. To reach 90% of the variance, 32 PCs must be considered, and 41 are required to reach 95%. When eigenvalues are very small as with the ones not considered here, it can be taken as an indication that the remaining variability is largely idiosyncratic and carries little trend. If allometry were strongly present (*i.e.* a major aspect of shape variation were size-dependent), some aspect of it should be visible in one or a set of eigenvectors among those that capture most of the shape information. It should therefore be possible to

construct a linear model of these ten PCs that can "explain" a large proportion of the variance in centroid size. While it is possible that an allometric signal could remain in the eigenvectors not considered here, any failure to appear in these ten would constitute reasonable evidence that any such allometric factor is small and disorganized, and is not a major factor in the shapes of adult *Pan* basicrania.

Multiple Regression analysis for allometry

To test the possibility that chimpanzee and bonobo basicrania exhibit a strong and consistent pattern of allometry, centroid size was regressed against the first ten PCs with Stata 8.2 (StataCorp, 2003). If allometry is a factor in basicranial shape in *Pan*, it should be possible to construct linear models relating the PCs to centroid size. As noted above, these ten PCs encompass most of the variance in the Procrustes residuals, and subsequent PCs account for very small proportions of the variance. The backward stepwise model selection approach was chosen for this task, because a multivariate approach to allometry is highly preferable. This approach begins with a model based on all predictor variables. It sequentially removes predictors that do not significantly contribute to the overall model's effectiveness, and can reinsert any that become significant again after confounding predictors are removed (Neter et al., 1996). The default situation for this approach is retention of predictor variables (in this case, principal components) in the model, corresponding to a rejection of the null hypothesis of no allometric relationship. The simultaneous consideration of several shape variables allows shape to be characterized more accurately, and makes it more likely that a relationship between size

and shape can be found. The backward stepwise regression approach is therefore the most conservative way to check for a size-shape relationship.

Results

The results of a series of regression analyses on these ten principal components and centroid size, including the primary backwards-stepwise run, are presented below.

Species distinctions in basicranial shape

Figure 3.2 below shows the first principal component values for the *P. paniscus* sample and the three subspecies samples of *P. troglodytes*. In this sample, the first principal component has achieved a perfect separation between species. Despite considerable overlap in centroid sizes, all specimens with PC1 values over 0.5 (the dotted line in the figure) are bonobos, and all others are chimpanzees[15]. Importantly, this result also indicates that 3D shape variation in the basicranium can be used to describe species-level variation in at least some hominoids. While it is certainly possible that larger samples would yield sufficient specimens with intermediate values to induce an overlap, it is clear that chimpanzees and bonobos have distinct basicranial shapes, and PC1 offers a reasonable representation of that spectrum. It is easy to see, given this distinction in shape, how anatomists such as Coolidge (1933) could recognize a species-level distinction between bonobos and chimpanzees. The fact that a reliably species-separating PC is the first indicates that species membership figures prominently in the pattern of

[15] The largest chimpanzee PC1 value is 0.49861; the smallest bonobo value is 0.54153.

basicranial variation in *Pan*. This finding offers strong support for the use of basicranial morphometrics in taxonomic studies such as this one. From Appendix 2, we can see that this PC alone accounts for nearly 17% of overall shape variation. It would be possible to construct a discriminant function that could predict *Pan* species membership with reasonable accuracy, though that is beyond the scope of this project.

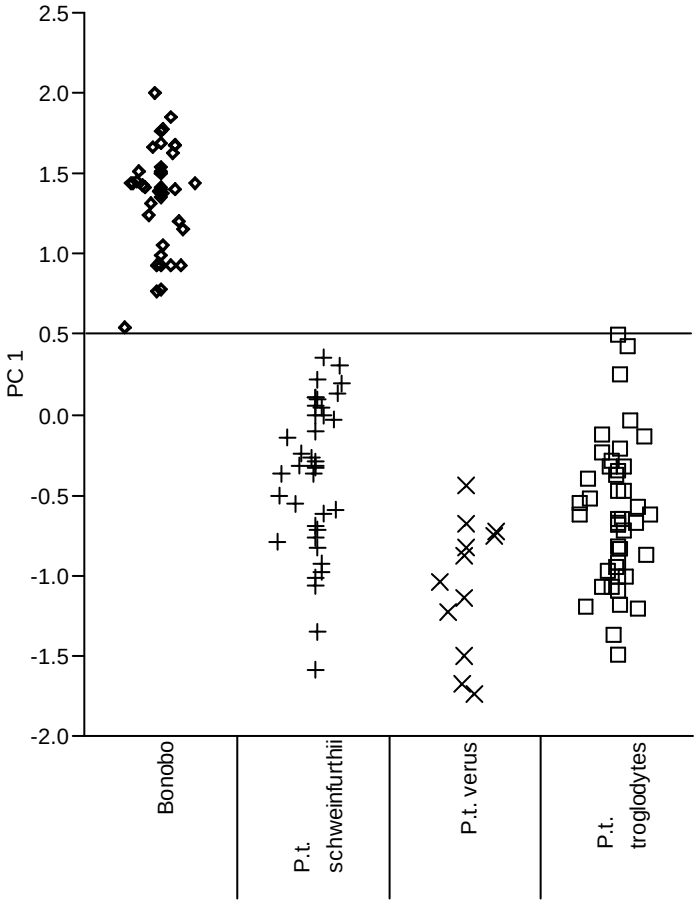

Figure 3.2. PC1 vs. centroid size for overall *Pan* sample.

It is possible to examine the "loadings" of the original variables on PC1 scores in an effort to understand the basic differences between chimpanzee and bonobo basicranial

shape. For example, relative to chimpanzees, bonobos tend to have anteriorly placed postglenoid processes, external auditory meati, and styloid pits, and laterally placed stylomastoid foramina. Chimpanzees tend to have anteriorly placed carotid canals and foramina ovale and spinosum, and their mastoid processes are placed much more laterally. To illustrate this distinction, *Morphologika2* (O'Higgins and Jones, 2006) was used to create a wireframe diagram of the basic basicranial anatomy (see Figure 3.3) and then to generate hyper-chimpanzee and hyper-bonobo shapes as shown in Figure 3.4 below.

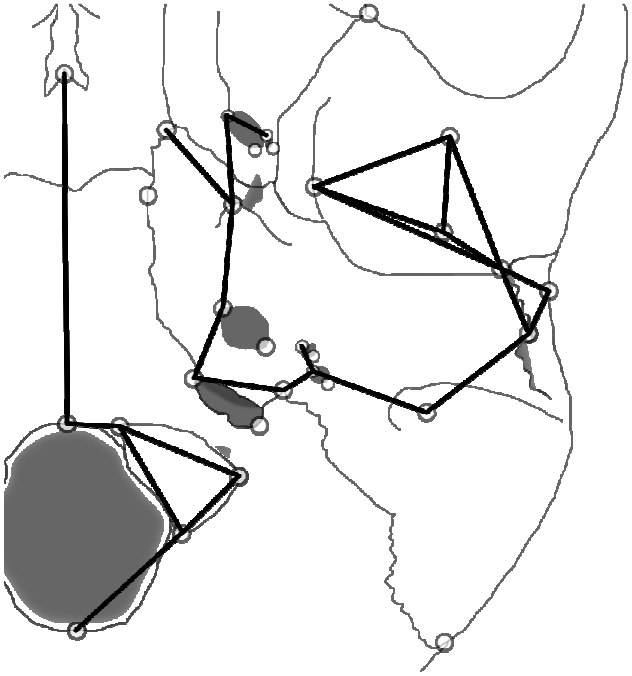

Figure 3.3. Wireframe diagram used to evaluate species distinctions in morphometric landmark configuration. See Figure 2.1 and Table 2.3 for landmark definitions.

While there is a rough *Pan*-wide expression of a size-shape relationship (chimpanzees tend to be larger than bonobos, and each species has a distinct shape, so size and shape are not entirely independent in the genus as a whole), it does not

automatically follow that this relationship must be present within the species themselves. This issue is examined below.

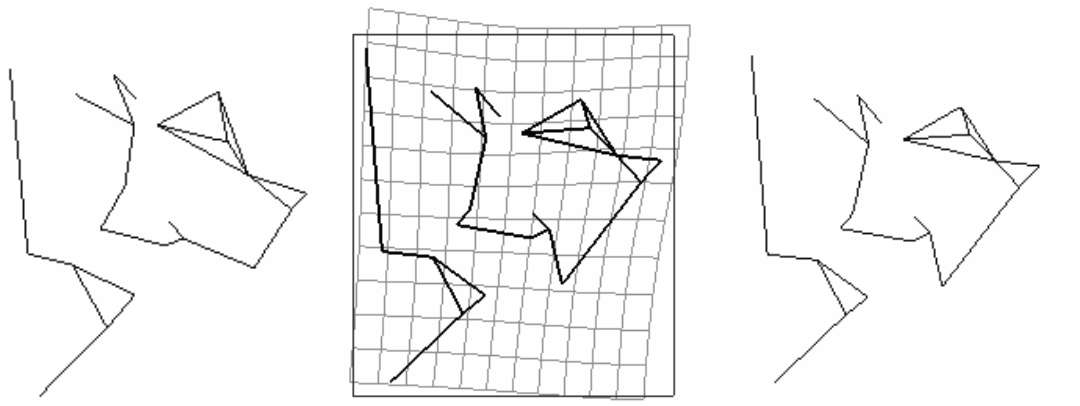

Figure 3.4. Hyper-chimpanzee (left) and hyper-bonobo (right) basicranial shapes in the transverse plane. The middle part of the figure is the thin-plate spline superimposition of bonobo (target) on the chimpanzee (reference) shape.

Existence of basicranial allometry within chimpanzees and bonobos

For bonobos, no combination of the first ten PCs yields a significant centroid size vs. PCs model. In other words, an allometric relationship with size is not visible in the first 60% of their shape variance. For chimpanzees, PC 3 alone (6% of overall shape variation) does provide a statistically significant model, but this model's r^2 value is only 0.07 ($p = 0.01$). The pattern of allometry in chimpanzee and bonobo basicrania is therefore not consistent. It is absent in bonobos, and present but weak in chimpanzees.

These results are in stark contrast with, for example, Singleton's (2002) results indicating that a shape PC capturing 67% of total cranial shape variation in adult

papionins correlated very strongly ($r^2 = 0.92$) with logged centroid size.[16] Singleton, however, examined the entire cranium, which includes skeletal structures that support dentition and surround the brain. The relative size of the dentition depends largely on sex, and both sex and brain size are strongly associated with body size differences. One would therefore expect a much stronger size-shape relationship among that dataset. The results here also contrast with McNulty's (2004), in which the first shape PC, comprising almost 47% of the shape variance of extant hominoid crania, correlated with centroid size at an r^2 of 0.88. In those analyses, PCs which described a relatively large portion of shape variation correlated very strongly with size, and it was clearly appropriate to perform a size-based shape correction before proceeding with the analysis. Here, a shape PC that describes only 6% of shape variation correlates weakly with size ($r^2 = 0.07$). It is not clear that an attempt to perform a size-based shape correction is appropriate in this case.

As the basicranium is subject to different developmental pressures from the cranium as a whole (Lieberman et al., 2000a), it is not entirely surprising that the degree to which allometry is expressed here may vary from that of the cranium overall. Other researchers have also found reduced or absent allometry in relatively restricted anatomical regions. Niewoehner (2005), for example, found no significant correlation between size and shape in a sample of Late Pleistocene human first metacarpal bases, and Lockwood (1997) found that infraorbital morphology was not correlated with craniofacial size in *A. africanus*.

[16] While it is often the case that the first PC represents size, Singleton's analysis also involved GPA, so her PC represents shape variation, as GPA involves an attempt at the removal of raw size.

Sex distinctions in basicranial shape

Given the success of PC1 at distinguishing species, it stands to reason that there could be one or several that distinguish specimens by sex. This information could hypothetically be used to diagnose sex among the fossils, and then to set the demography of the reference samples of *Pan* crania. To test this possibility, specimens' PC scores were compared by sex, using a series of t-tests. None of the shape PCs distinguish specimens by sex in the pooled sample. It would therefore appear that while this set of landmarks can distinguish species, it cannot distinguish sexes. To further confirm this result, though, the sexes' PC scores were compared in separate-species samples.

While the pooled sample does not do so, the separate species samples independently have some PCs that tend to correspond to sex. When corrections for multiple comparisons (Wright, 1992) are not considered, bonobos are significantly different by sex for PCs 1 (the species-separating PC; male bonobos appear to be more chimp-like and females more bonobo-like[17]) and 7 (p = 0.0145 and 0.0435, respectively), but there is substantial overlap in the sexes' ranges. The chimpanzee sample shows distinctions by sex for PCs 3 and 4 (p = 0.0358 and 0.0117, respectively), also with substantial overlap.

This set of comparisons, however, involved a battery of 20 separate tests (two species, ten PCs)[18]. The Bonferroni-adjusted critical p value to achieve an overall

[17] This result is quite interesting, given the absence of a statistically significant tendency for larger bonobos to be more chimp-like. It may be inferred that larger females are not more chimp-like, and smaller males are not more bonobo-like.

[18] It could be argued that the tests for the pooled-species sample should also be considered, making a total of 30 separate (but not entirely independent) comparisons, but the critical p values based on 20 tests reported above are less conservative and therefore more likely to show any support for an inference of significance. In any event, none of these tests were significant at Bonferroni- or Šidák-adjusted overall α = 0.05 for 20 tests, and definitely would not be for 30.

experimentwise α of 0.05 in this situation is 0.0025, and the Šidák-adjusted value is 0.00256 (Wright, 1992). None of the above *p* values approach these critical values, however, so one cannot say with 95% confidence that these PCs accurately capture sex distinctions within the two species.

The third principal component (PC3) has therefore made an appearance in the analyses of allometry and of sexual dimorphism, but only for common chimpanzees. It seems to play some role in the slight expression of static allometry among them, as well as in their sexual dimorphism. The third PC thus encompasses some degree of sexual dimorphism in both size and shape for chimpanzees, but not bonobos. The species have independent patterns of basicranial size and shape variation.

It may have been possible to use a consistent pattern of sexual dimorphism in the *Pan* sample to diagnose sex in the fossil specimens, and then to use this pattern in an attempt to refine the demographics of the reference *Pan* samples for the tests of fossil variability that form the heart of this project. There is not, however, a PC within the first ten (encompassing 60% of the total variation) that can distinguish sex in the pooled sample, which is the situation that would have to be applied to the fossils. The few PCs that seem to vary by sex in the separate-species samples do not withstand multiple-comparisons adjustments, and are not consistent between species. Therefore, the shapes of *Pan* basicrania as described by these landmarks do not provide a means for diagnosing sex in the fossil specimens, and we cannot look to sexual dimorphism in modern *Pan* for guidance in using the landmark configurations of the fossils to help select the demographics of the model *Pan* samples (let alone apply information gained from those fossil-informed model samples back to the fossils). This inference is consistent with the

findings of Cobb and O'Higgins (2004), who argue that these shape trajectories are not necessarily different by sex.

Conclusions

Two species closely related to hominins have been examined here for the possibility of providing a model by which fossil shape could be "corrected" for size, and neither one offers such a model. Adult chimpanzees and bonobos do not share a genus-wide pattern of size-shape covariation. Adult bonobos do not exhibit an allometric pattern that is visible within 60% of the shape variation that they share in common with chimpanzees. Chimps do exhibit such a pattern, but it is quite weak. Neither of these species therefore offers a trajectory through shape-space that could be extrapolated to construct a size-based shape adjustment for the fossils. Others have made this type of adjustment to morphometric datasets, but always in cases in which the size-shape relationship was much stronger. Given the differences between the allometric signals of chimpanzees and bonobos, it would not be appropriate to force an allometric correction on the fossil data. It is reasonable to proceed on the data "as-is," with the caveat that this relationship was only tested for adult specimens and so is relevant to static allometry only, as opposed to ontogenetic growth trajectories. It must be noted, however, that allometry is not entirely absent in at least chimpanzees' basicrania. It should also be noted that growth trajectories (as distinct from static adult allometry) may be sufficiently similar in modern hominines to allow the use of extant growth trajectories to estimate the

adult shape of juvenile fossils (McNulty et al., 2006), but the data considered here involve only adult chimpanzee and bonobo specimens.

While there are within-species sex differences in shape at first glance, there is not a shape difference between sexes that is common to the two species. Further, a multiple-comparisons adjustment makes the within-species sex differences statistically nonsignificant. Thus, there is not a consistent pattern of sexual dimorphism in the basicranial landmark configurations in the *Pan* sample. Had such a pattern existed, it would have been useful in attempts to diagnose sex in the fossil sample, and to use the resulting sex ratios for guidance in constructing *Pan* reference samples with appropriate demographics for the variability comparisons in the chapter(s) to follow. Without such guidance, the most advisable (or alternatively, least objectionable) route is to use reference samples with equal membership by both sex and species.

Chapter 4. Results

As described in the Methods and Materials chapter, a total of 21 tests were performed in which inter-fossil pairwise Procrustes distances were compared to the distribution of inter-*Pan* pairwise Procrustes distances (see Appendix 1). The results of these individual tests were used to generate a collection of interspecimen distances defined as the average percentile of the fossils' Procrustes distances in the corresponding distributions of all-*Pan* pairwise Procrustes distances, with that average weighted by the number of landmarks per test. This nonmetric operationalization of distance (*i.e.* the shape difference between the 50^{th} and 51^{st} percentiles is likely to be much smaller than that between the 90^{th} and 91^{st}), effectively captures the relative degree of difference between fossil specimens as compared to the distribution of *Pan* pairwise distances (see "Advantages of using an indirect distance proxy" section below). The weighted averages are presented in Figure 4.1.

This approach is intended to deal with the tradeoff between numbers of specimens and numbers of landmarks. As one considers more specimens, the number of landmarks common to each decreases. Likewise, increasing the number of landmarks involved in a comparison reduces the number of available fossils. For each of the ordinations discussed below, the approach used here allowed specimens with numerous common landmarks to be compared with the most thorough dataset available, while including other less-complete specimens in the same ordination.

It can be seen in Figure 4.1 that the specimen included in the most pairs with percentiles beyond 95 (the dashed line) is Sts 19, with Stw 580, Stw 187, and Sts 5 close behind. The large Procrustes distances involving Sts 19 are unsurprising, as this

specimen has already been shown to be morphologically atypical among *A. africanus* specimens, though the taxonomic importance of these differences is controversial (Ahern, 1998; Kimbel and Rak, 1993). Unfortunately, all of these pairs involve tests with low numbers of landmarks ($k \leq 7$), so it is possible that they are not particularly diagnostic. A specimen involved in comparisons with larger numbers of tests and more landmarks is Stw 505, and these comparisons also tend to be associated with high percentile scores. While insufficient by themselves to warrant splitting *A. africanus*, these results indicate that some specimens have different shapes than others traditionally attributed to *A. africanus*, and merit further attention.

Ordination of morphometric distances

Ordination of morphometric distances provides a way to visualize patterns and relationships, and is an ideal way in which to assess a dataset for the presence of specimen groups that could correspond to taxonomic distinctions. The distance observations described above are ordinated here with both principal coordinates analysis and nonmetric multidimensional scaling.

Principal Coordinates Analysis

An overall ordination of all specimen pairs' percentile scores was not possible, as many pairs of fossils had no landmarks in common (e.g. Sts 71 with each of Stw 266, Stw 580, Stw 187, and Stw 329) or very few (e.g. Stw 505 with each of MLD 31, Stw 580, and Stw 187). Ordinations involving smaller sets of fossils, however, are possible because each pair of fossils in these smaller sets were included in at least one test.

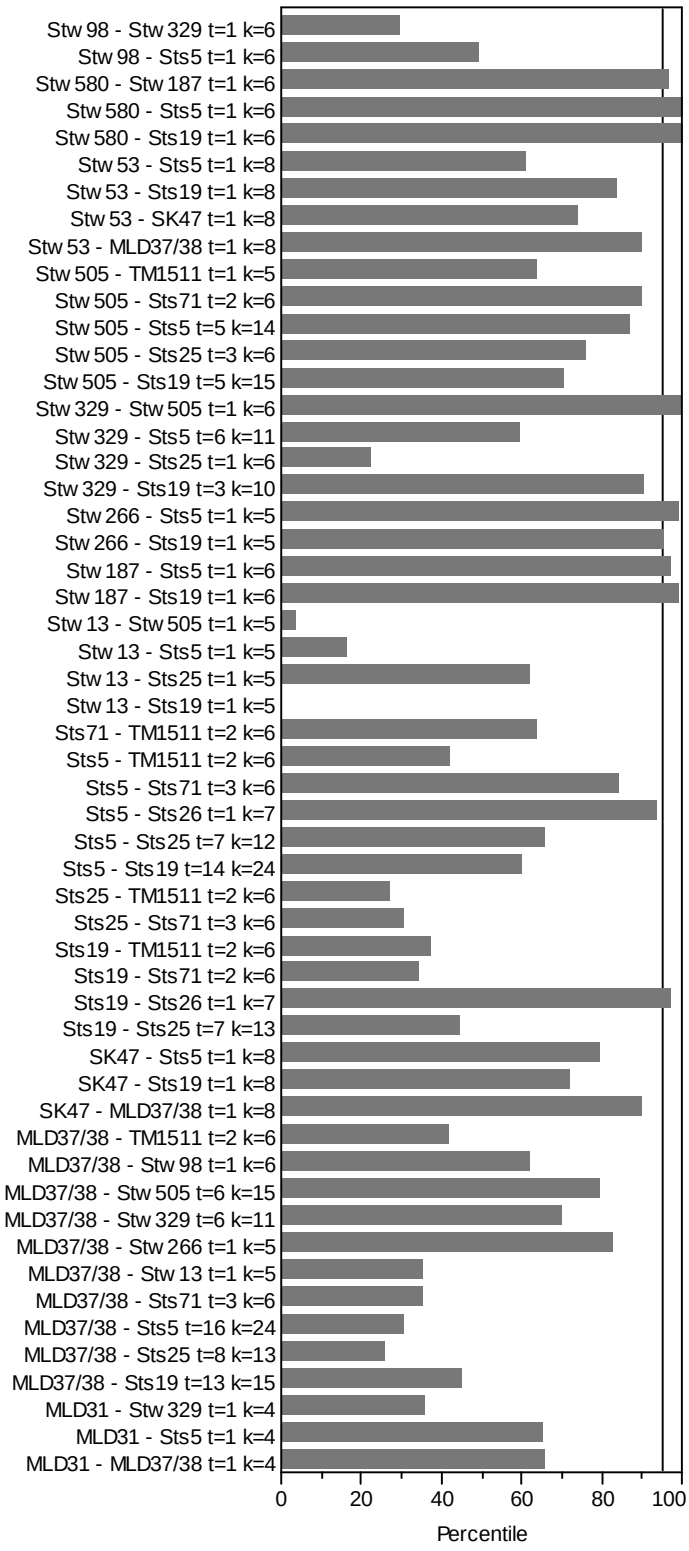

Figure 4.1. Weighted average percentile rank of fossil pairs' morphometric Procrustes distances in all-Pan distribution of pairwise Procrustes distances. Dashed line = 95th percentile, t = number of tests including a given pair of fossils, k = maximum number of landmarks in a relevant test.

A “core” group of specimens with tests in common was MLD 37/38, Sts 5, Sts 19, Stw 505, and Sts 25. Some other specimens maintain landmarks in common with these five, but not necessarily with one another. In order to overcome this difficulty, three sets were constructed in which these other specimens were individually added to the core group. The first set eligible for ordination involved the core group fossils as well as TM 1511 and Sts 71. The second set included the core group and Stw 329, and the third the core group with Stw 13.

In the principal coordinates (PCO) analyses here, the data were power transformed by power c = 1, 2, 4, or 6. This step can be viewed as a detrending operation, or as a means of modifying (hopefully reducing) the extent of non-Euclidean relationships between specimens. Higher values of c tend to concentrate more variance in the first coordinate axis. This concentration may simplify visualization of the data, but can also create negative eigenvalues for other axes, which would make the projections non-Euclidean (Cailliez and Pagès, 1976; Podani and Miklós, 2002). If large enough, these departures from the Euclidean condition could make the 2D or 3D projections of the ordinations difficult or impossible to interpret. To minimize this possibility, the ordinations for all three sets of fossils were tried with all four available values of c, and the value that minimized negative eigenvalues was used. This transformation would not have an effect on the identification of divergent specimens which is the goal of the ordinations, because it cannot modify the relative distances between specimens. The results are shown in Table 4.1.

Fossil set	*c*	Eigenvalue results for coordinate axes
MLD 37/38, Sts 5, Sts 19, Sts 25, Stw 505, TM 1511, Sts 71	1	Eigenvalues for all axes are negative
	2	Four positive eigenvalues, one near zero, one small negative[19]
	4	Three positive eigenvalues, one near zero, two negative (one large)
	6	Three positive eigenvalues, one near zero, two large negative
MLD 37/38, Sts 5, Sts 19, Sts 25, Stw 505, Stw 329	1	One positive eigenvalue, one near zero, three large negative
	2	Four positive eigenvalues, one near zero
	4	Three positive eigenvalues, one near zero, one small negative
	6	Two positive eigenvalues, one near zero, two small negative
MLD 37/38, Sts 5, Sts 19, Sts 25, Stw 505, Stw 13	1	One positive eigenvalue, one near zero, three large negative
	2	Three positive eigenvalues, one near zero, one small negative
	4	Three positive eigenvalues, one near zero, one large negative
	6	Two positive eigenvalues, one near zero, two small negative

Table 4.1. Effects of choice of *c* values on ordination of fossil distances for morphometric data. See footnote for discussion of *small* vs. *large* negative eigenvectors.

Table 4.1 shows that the choice of *c* value is not trivial, consistent with the arguments of Podani and Miklós (2002). Using $c = 1$ would be inappropriate because it returns mostly (or even all) negative eigenvalues. This indicates that the transformed distance matrix is completely non-Euclidean, and that any effort to visualize it as such would be imperfect at best (Cailliez and Pagès, 1976). The choices of $c = 4$ or 6, while providing large eigenvalues on the first axes (data not shown) that would make for easy visualization in either 2D or 3D plots, both result in at least one large negative eigenvalue for one or more of the tests. Therefore, in order to maintain the best approximation of Euclidean-ness for the ordinated specimen relationships, the *c* value of 2 was applied to all three ordinations. The presence of two small negative eigenvalues (one each for the first and third fossil sets) should not present a problem, as the absolute value of each is less than half that of the smallest positive eigenvalue in the corresponding ordination, as

[19] A negative eigenvalue is described here as *small* if its absolute value is less than that of the smallest positive one, and *large* if greater, as suggested by Digby and Kempton (1987; cited and discussed in Podani and Miklós, 2002). The ones described as *near zero* have absolute values smaller than 10^{-13}.

shown in Table 4.2 (Cailliez and Pagès, 1976; also Digby and Kempton, 1987; cited in Podani and Miklós, 2002). While substantial deviations from a purely Euclidean situation should be avoided, minor deviations are acceptable (Manly, 1994). The small negative eigenvalues obtained here with $c = 2$ should therefore present little obstacle to visualization.

The $c = 2$ power transformation in PCO as described above produced eigenvalue results as shown in Table 4.2. The "Sum Percents" column is an approximate index of agreement between the ordinated relationships and those in the original data, similar to a cophenetic correlation.

Fossil set	Axis	Eigenvalue	Percent	Sum Percents
MLD 37/38, Sts 5, Sts 19, Sts 25, Stw 505, TM 1511, Sts 71	1	5013.2	40.5 %	87.2 %
	2	4036.2	32.6	
	3	1160.1	9.4	
	4	996.8	8.1	
	5	near 0	0	
	6	-460.7	-3.4	
MLD 37/38, Sts 5, Sts 19, Sts 25, Stw 505, Stw 329	1	6033.3	44.8 %	90.3
	2	3438.3	25.6	
	3	2144.0	15.9	
	4	536.6	4.0	
	5	near 0	0	
MLD 37/38, Sts 5, Sts 19, Sts 25, Stw 505, Stw 13	1	4922.0	45.9%	81.7
	2	2788.6	26.0	
	3	1267.2	11.8	
	4	near 0	0	
	5	-219.6	-2.0	

Table 4.2. Eigenvalues and corresponding percentages for coordinate axes following application of PCO to morphometric data, with transformation coefficient $c = 2$, on three sets of fossils. Summed percentages include negative eigenvalues.

A three-dimensional approach was selected for plotting the PCO ordination results. The first three coordinate axes for the fossil sets encompass 82.5, 86.3, and 83.7% of the overall variance in the data for each of the ordinations respectively (though offset somewhat by the presence of small negative eigenvalues for the first and third ordinations). Two-dimensional plotting (*i.e.* disregarding the third axis) would result in loss of 9-16% of the total variance relative to the 3D option.

PCO ordinations

These ordinations are plotted in Figures 4.2 (for the first set of specimens listed above), 4.3 (second set), and 4.4 (third set). The plots show the specimens' PCO scores for the three ordinations. The marker size for each specimen corresponds to its maximum number of landmarks included in relevant tests. The relationships of specimens indicated by larger markers, therefore, are likely to be more reliable. Relative placement of those indicated by smaller markers is more contingent, and may be affected by sampling bias created by reduced preservation.

The three PCO ordinations yielded several consistent results. First, Sts 5 and MLD 37/38 appear to be quite similar in these analyses, as they are consistently ordinated near one another. As indicated by the large size of their markers, this result is based on a large number of morphometric landmarks. For its part, the large Stw 505 specimen ordinates in distinction from the others, and this difference is usually most pronounced on the first axis, which encompasses 41 – 46% of the overall variation in the three datasets (Table 4.2). Another consistent result is that Sts 19 and Sts 25 ordinate at the margins of

the specimen distribution, but always along minor axes that capture 9 – 26% of the overall variation.

Specimens that appear in only one ordination each (i.e. not members of the "core group") also show some separation from the others. This is most notable with Stw 329 in Figure 4.3 and Stw 13 in Figure 4.4, but also visible with Sts 71 in Figure 4.2. The separation of Stw 13 and Stw 329 lies along the first axis, as with Stw 505. Sts 71, on the other hand, appears to resemble Sts 19 to some extent. The importance of these patterns is explored below.

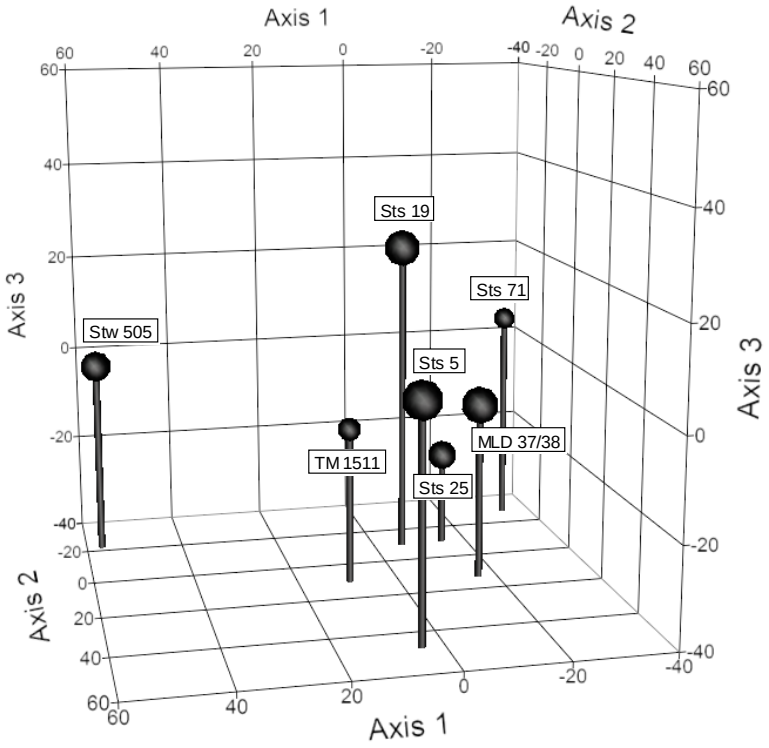

Figure 4.2. PCO ordination of distances from morphometric data for the first set of fossil specimens. Marker size corresponds to maximum number of landmarks used in a given test.

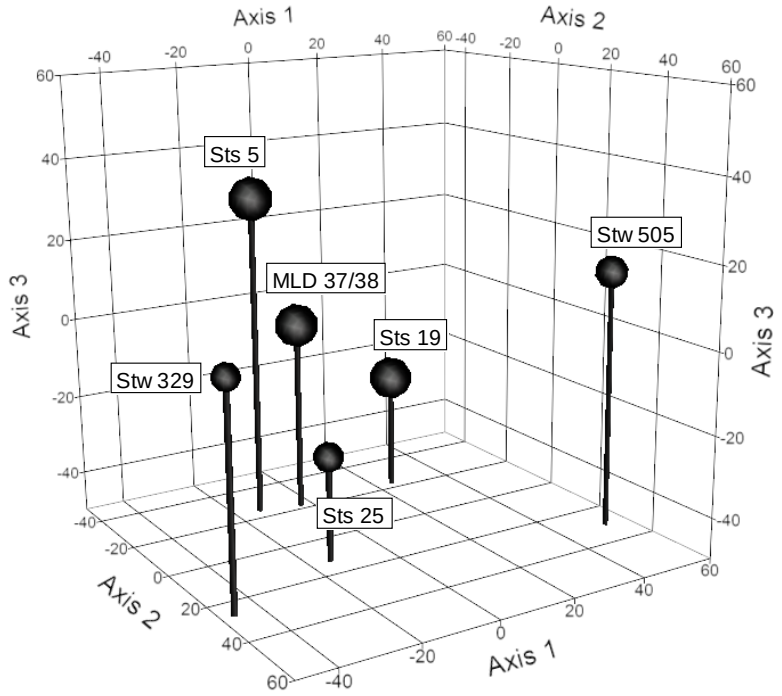

Figure 4.3. PCO ordination of distances from morphometric data for the second set of specimens. Marker size as in Figure 4.2.

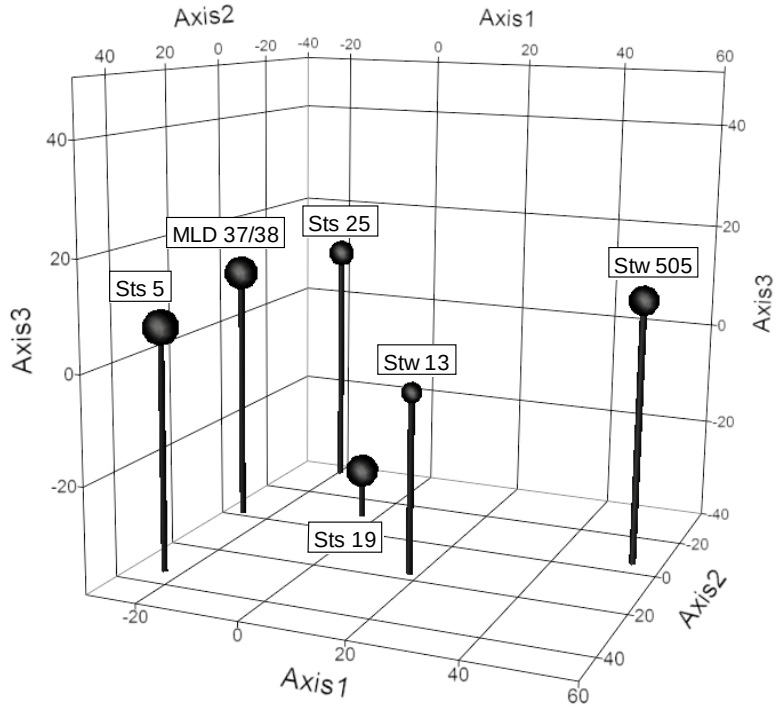

Figure 4.4. PCO ordination of distances from morphometric data for the third set of specimens. Marker size as in Figure 4.2.

Nonmetric Multidimensional Scaling

In order to further evaluate shape variation in the sample, repeated nonmetric multidimensional scaling (NMDS) runs were performed on the same matrices of interspecimen differences as used in the PCO analyses (see Chapter 2 for discussion). For the first group (the core set of five specimens; MLD 37/38, Sts 5, Sts 19, Sts 25, and Stw 505; plus TM 1511 and Sts 71) and second group (the core set plus Stw 329) the NMDS analyses were repeated 100 times.

On the first set, a minimum stress value of 0.03476 was obtained in five of the 100 runs, and Procrustes superimposition of the resulting configurations of specimens in *PAST* (to remove reflections) followed by *tps-Small* indicated that these five configurations were essentially identical, allowing for rotations (Procrustes distances between configurations ranged from 0.01 to 0.03). For the second set, the minimum stress value was 0.04016, obtained in ten runs of the 100. The same Procrustes superimposition procedure indicated that these ten configurations were essentially identical as well.

For the third group of specimens, the runs quickly converged on an optimal 3-D configuration with zero stress (this value was obtained in nine of the first ten runs) with identical coordinates for each of the six specimens, so the analysis was repeated no further. It must be emphasized that these stress values do *not* represent a p-value for a test of any hypothesis. They simply reflect the degree to which the depicted interspecimen ranked distances depart from the original-data ranked differences. For n = 6 or 7 specimens and m = 3 dimensions, ordination stresses meeting Kruskal's (1964a)

criteria of “good” (stress ≤ 0.05) or “excellent” (stress ≤ 0.025) are the rule, not the exception (Klahr, 1969).

NMDS ordinations

The NMDS ordinations are presented in Figures 4.5-4.7 below. As with the PCO plots, the size of the dots corresponds to the maximum number of landmarks included in a relevant test. Note that as the NMDS algorithm seeks to match the sequence of ranked distances in projection-space (what is shown here) to the ranks of differences in the original data, the projected distance between specimens matters less than their placement within or at the margin of a distribution.

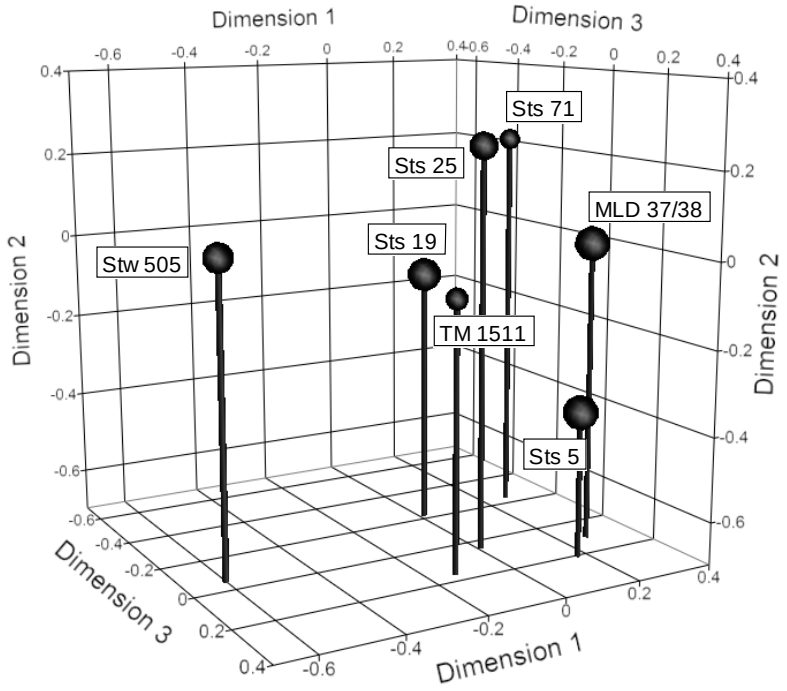

Figure 4.5. NMDS ordination of distances from morphometric data for first set of specimens; stress = 0.03476. Marker size corresponds to maximum number of landmarks used in a given test.

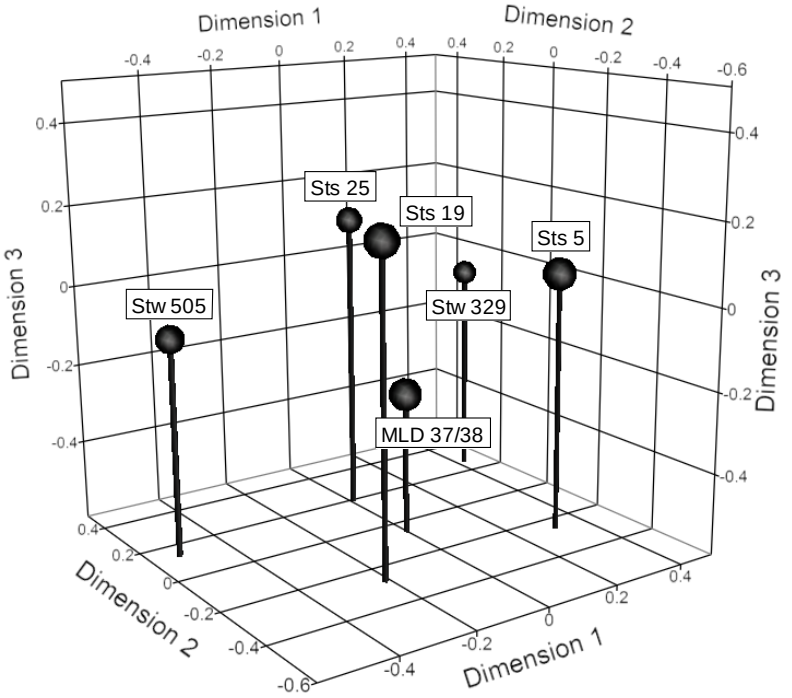

Figure 4.6. NMDS ordination of distances from morphometric data for second set of specimens; stress = 0.04016. Marker size as in Figure 4.5.

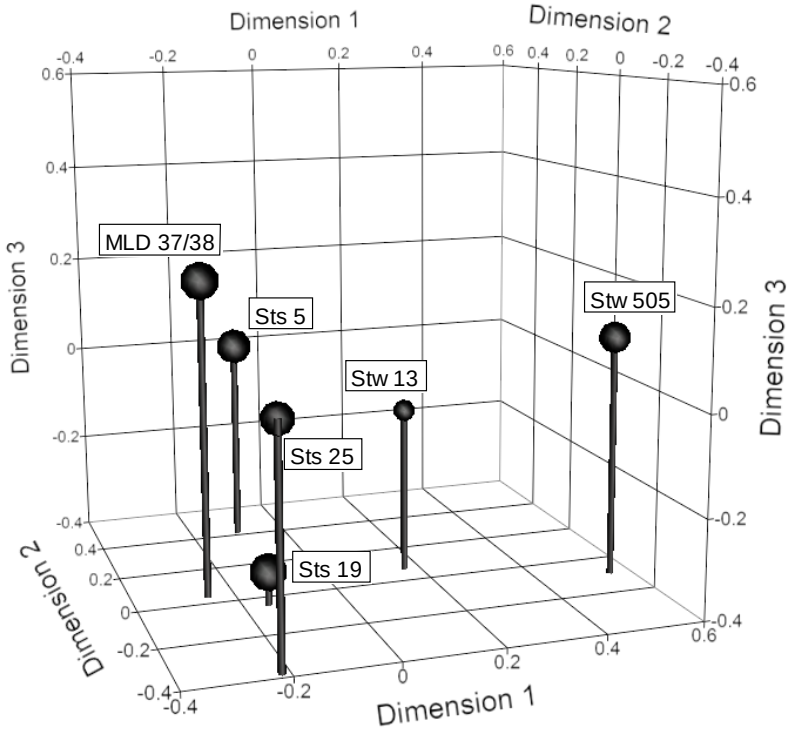

Figure 4.7. NMDS ordination of distances from morphometric data for third set of specimens; stress = zero. Marker size as in Figure 4.5.

In comparison of Figures 4.5-4.7 with Figures 4.2-4.4, it can be seen that the primary results are generally consistent throughout, despite the differences in approaches to ordinating these data. The Sts 5 and MLD 37/38 fossils' basicrania are similar, although the similarity of Stw 329 to each of them as shown in Figure 4.6 tends to separate them somewhat[20]. Stw 505 remains distinct from the others. Sts 19 and Sts 25 consistently ordinate at the margins of the main body of *A. africanus* specimens (excluding Stw 505). Stw 329 again ordinates away from the others, and in a direction opposite that of Stw 505[21]; Stw 13 is midway between Stw 505 and the others; and Sts 71 also ordinates near the margin of the distribution.

Summary of morphometric data ordination results

In summary of the overall ordination results, Stw 505 appears to remain distinct from the other specimens, with Stw 13 midway along the axis separating it from the main cluster. Sts 19 *may* also be distinct, and Sts 71 may join it, but in a direction orthogonal to the one separating Stw 505 from the others. It also appears that Sts 25 is consistently located at the margins of the distribution, but it does not have such a clear relationship with either Stw 505 or Sts 19. It can thus be hypothesized based on the ordination results that Stw 505 is the most likely to belong to a taxon distinct from *A. africanus*, but Sts 19, Sts 71, and Sts 25 also give some indication of separation. It does not appear to be likely, based on these results, that Stw 505 and the possible Sts 19 / Sts 71 pair are both distinct from *A. africanus*, and also members of the same taxon. Because Sts 25 is not a member

[20] NMDS ordinations project ranked distances, so when Specimen A is even slightly more similar to Specimens B and C than they are to one another, the B – C distance can appear to be inflated.
[21] Unlike the eigenvector-based approaches of PCA and PCO, NMDS produces axes that are entirely arbitrary.

of such a consistent relationship, it is difficult to make such a pronouncement about it, but it may still be distinct from the bulk of the specimens. The disposition of Stw 13 is also of interest, due to the consistency of its location midway between Stw 505 and the main cluster. Its morphological traits will bear heavily on the question of whether it aligns more closely with Stw 505 or with the main cluster, or serves to connect them.

Advantages of using an indirect distance proxy

As described in Chapter 2, the pairwise distances between fossils were expressed indirectly for the ordinations. Rather than the original Procrustes distances, the metric used here was the percentile rank of that Procrustes distance in the overall *Pan* dataset, averaged with weight given to the number of landmarks in a given test. For the first ordination group, specimens had as many as 24 landmarks in common (Sts 5 with both MLD 37/38 and Sts 19) and as few as five (Stw 505 with TM 1511). In this ordination group, there is a set of five landmarks common to all specimens. Had Procrustes shape distances been used as the difference metric more directly, all pairwise distance observations would have been either limited to these five landmarks or subject to problems arising from the incommensurability of shape distances based on different numbers of landmarks. The ordination would then have been possible with direct distance observations, but much less informative. This is why the interspecimen distances were first referred to the distribution of *Pan* interspecimen distances.

The situation for the second ordination group was much different. Here, the lowest number of landmarks common to a fossil pair was six, for each of the pairwise comparisons among Stw 329, Stw 505, and Sts 25. The consistency of the number six

among these three fossils is, however, a mere coincidence. The three pairs do not involve the same six landmarks. Only one common landmark was used in all three of the tests comparing these pairs of specimens. This ordination would have been impossible without the indirect distance observation described above and in Chapter 2.

The third ordination group was much more straightforward. The smallest number of common landmarks in a test included in this ordination was five, for all of Stw 13's comparisons against the other specimens. This ordination would still have been possible with a more direct distance proxy, but it would have been limited to five landmarks for all fossil pairs, as for the first ordination group.

Specimen size and allometry

It was possible that an allometric signal in *Pan* basicrania could have been used to generate a model for performing a size-based shape correction to the fossil data (following McNulty, 2004). As shown in Chapter 3, however, the pattern of *Pan* basicranial shape variation did not lend itself to the construction of such a model for this dataset. As a result, all of the morphometric shape comparisons here are possibly confounded by shape variation that is directly associated with specimen size, as opposed to taxonomic differences. The most obvious specimen with morphometric differences from the remainder of the fossils is Stw 505 (see Figures 4.2-4.7). Others have also noted its large size (e.g. Lockwood and Tobias, 1999), and if allometry played much role in *A. africanus* basicranial shape (unlike bonobos and chimpanzees), this specimen would be the most strongly affected. It is also noted above that Stw 329 ordinates in some distinction from the other specimens, and along the same axis, but in the opposite

direction. Stw 505 is already known to be quite large. If this size is expressed in the datasets used here, and if Stw 329 is shown to have small size, this would provide an alternative explanation to explain the apparent distinctiveness of Stw 505 and Stw 329.

Table 4.3 below includes the centroid sizes for Stw 505 and Stw 329 in all morphometric comparison tests in which they were included. It also includes the centroid sizes for other specimens which appear to exhibit distinction from the others in these ordinations, as well as Sts 5 and MLD 37/38. It can be seen that Stw 505 is substantially larger than all other specimens in nearly all cases. The only two exceptions are shown in bold type, and they are limited to tests which involved five landmarks each. Test 15 was limited to landmarks around the foramen spinosum and glenoid region; test 24 used landmarks around the foramen ovale and glenoid. Stw 505's size is typical in this limited anatomical region, but it is larger than the others when more landmarks are considered. This pattern is likely responsible for Stw 505's shape distinction from the others.

Stw 329's position along the axis that separates Stw 505 from the other specimens becomes understandable when its centroid size is compared to those of the others. In all of the tests in which it was considered, Stw 329 is the smallest. It therefore appears that Axis 1 of Figure 4.3 and Dimension 1 of Figure 4.6 are associated with size variation, implying that the orthogonal axes/dimensions are more associated with shape variation. The fact that Sts 19 tends to separate from the others along these orthogonal axes, and its typical size, indicate that Sts 19's differences from the other specimens are not primarily allometric in nature.

Test	Landmark count	MLD 37/38	Sts 5	Sts 19	Sts 25	Sts 71	Stw 13	Stw 329	Stw 505
4	11	55.01	58.71					45.62	
5	6	35.95	37.68					30.45	
6	15	59.97		59.79					68.28
7	14	57.40	59.44	57.91					66.08
14	4	27.37	27.37					20.03	
15	5	19.94	23.63	21.98	19.88		23.68		**22.95**
20	6	37.38	40.58	38.24				33.42	47.60
21	10	50.36	53.20	49.66				43.24	
22	6	45.64	48.72	46.30	41.96			38.30	
23	6	36.86	40.62		34.44	36.47			42.79
24	5	24.11	27.16	25.39	21.79	21.80			**25.65**

Table 4.3. Centroid sizes of Stw 505 and for Stw 329 in all superimposition tests in which they were considered, as well as of selected other specimens (see text). Boldface type indicates departure from overall pattern of Stw 505 being larger than all other specimens and Stw 329 being smaller. See Appendix 1 for more details of these tests.

Index of variability

The other method employed to evaluate morphometric variation in *A. africanus* was the index of variability (Relethford and Blangero, 1990). As described in the Methods chapter, the results are expressed in terms of ratios of fossil variability to *Pan* variability, with each of the results including a given fossil averaged and weighted by the number of landmarks considered in each test. The ratios were logged for the averaging step, and transformed back to raw space for this chart. This approach gives an informal measure of the variability of *A. africanus* when a given fossil is included in it, subject to the sampling effects described earlier. The results of this procedure for each of the specimens included in the morphometric analyses are presented in Figure 4.8. The overall weighted mean of the logged ratios for all *A. africanus* tests was 0.035,

corresponding to a fossil : *Pan* variability ratio of 1.04. The fossil sample is slightly more variable than the *Pan* sample.

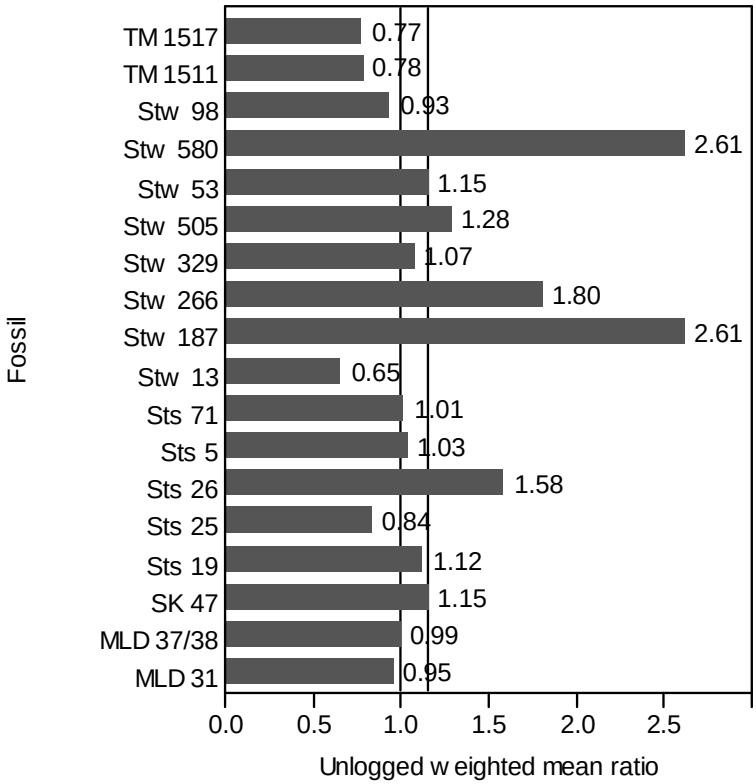

Figure 4.8. Unlogged weighted mean fossil : *Pan* variability ratios for fossil samples including listed specimens. Continuous dashed line (- - -): ratio 1.0; staggered dashed line (— - -): ratio 1.15. See text for discussion.

With this approach, one can see that variability of fossil subsets including Sts 5, MLD 37/38, and Sts 71 is approximately equal to *Pan* variability (their ratios are approximately 1). For sets including Sts 25 and/or Stw 13, fossil variability is somewhat less. Data are also available for three specimens that are viewed here as representing different species than *A. africanus*: Stw 53, which has been assigned to early *Homo* (Curnoe and Tobias, 2006), and SK 47 (de Ruiter et al., 2006) and TM 1517 (Broom and Schepers, 1946), which represent *A. robustus*. The unlogged weighted mean ratio for

both Stw 53 and SK 47 is 1.15 (fossil samples including these specimens are 15% more variable than *Pan*), and a staggered dashed line shows this level of variability in Figure 4.8. Curiously, the type specimen of *A. robustus*, TM 1517, is associated with a ratio of only 0.77. This specimen was compared against others with only five landmarks, mainly from the glenoid region of the temporal bone. It would appear that these few landmarks are insufficient to identify species in the fossil record. SK 47 and Stw 53, on the other hand, were tested on eight landmarks. These eight are mainly located on the petrosal, and appear to be better situated for identifying species-level variation.

Effect of sample size on variability scores

The sets involving Stw 580, Stw 266, and Stw 187 are substantially more variable than the corresponding *Pan* samples, and the set involving Stw 13 substantially less so. This is likely to represent some degree of sampling effect, as these extreme ratios are associated with tests involving small numbers of landmarks and/or specimens. See Figure 4.9. This figure returns to the logged variability ratios for ease of visualization. It can be seen that specimens with relatively few available landmarks (maximum *k* [□ icons] less than 8) could only be included in 3 or fewer tests [× icons] each. This category includes specimens at all points in the range of logged ratio values: roughly -0.5 to 1, indicating fossil : *Pan* variability ratios of 0.61 to 2.7 (ratios higher than 1 indicating greater fossil variability). As the number of landmarks and therefore tests increases, the range of logged ratios decreases: first to about -0.2 to 0.25 (corresponding to raw variability ratios of 0.82 to 1.28), and finally to roughly 0 to 0.1 (raw ratios of 1 to 1.1). While extreme values of this ratio are interesting, it is possible that they reflect sampling

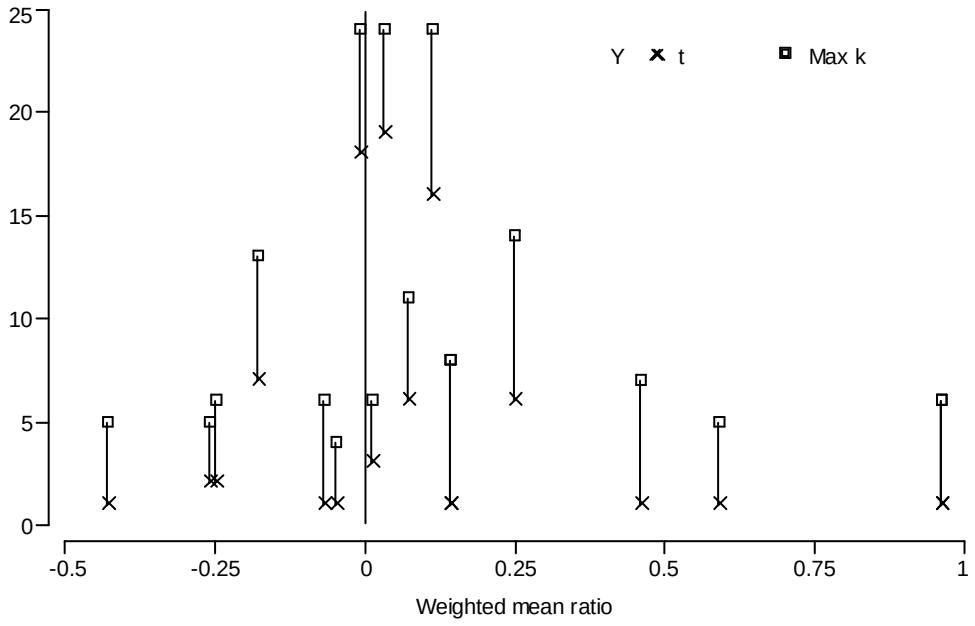

Figure 4.9. Relationships between number of tests, maximum number of landmarks (*k*) in those tests, and weighted mean logged variability ratios. Lines connect markers for individual specimens. Vertical dotted line: logged variability ratio = 0, indicating equal variability in fossil and *Pan* samples.

effects due to low numbers of both landmarks and tests rather than taxonomically-important variation[22]. The sample of well-preserved specimens is small, but tends to indicate that extreme departures from the scale of variability expressed by the two species of *Pan* are unlikely in *A. africanus* and any secondary species that may be subsumed within it. Rather, it appears that this sample of specimens is, overall, slightly more variable than *Pan*. This finding is interesting, but by itself is not dispositive of the question of whether multiple species are subsumed within *A. africanus*. It serves, rather, to support the finding that some specimens' variation is inconsistent with the hypothesis of monotaxy. Had the fossils not shown more variation than a modern analog (see

[22] Alternatively, it is possible that the small number of relatively complete specimens could be causing its own sampling effect and artificially deflating apparent variation.

Chapters 1 and 5 for discussion of the appropriateness of the one chosen here), it would be difficult to justify any conclusion that the fossils contain more than one species.

Key specimens identified with morphometric techniques

The method of comparing variability scores may be less reliable than that of comparing Procrustes distances. Nevertheless, this technique points to specimens Stw 505, Sts 26, Stw 266, Stw 580, Stw 187, Sts 19, and Stw 329 as being included in samples with more variation than is seen in *Pan*, and all but the latter two associated with greater variation than the fossil sample including *A. africanus* as well as a few specimens of early *Homo* and *A. robustus*. The ordinations described above suggest that Stw 505 is morphometrically distinctive from other *A. africanus* specimens, and that Sts 19, Sts 25, Sts 71, and Stw 13 are as well, although to a lesser extent. Finally, the raw un-ordinated Procrustes distances shown in Figure 4.1 indicated that Sts 19, Sts 26, Stw 505, Stw 580, and Stw 187 tended to be involved in distances at or above the 95th percentile of corresponding pairwise *Pan* distances. The specimens indicated by all three approaches as being at least somewhat distinct are Stw 505 and Sts 19. As noted above, however, Stw 505 is unusually large on almost all landmark configurations considered here. It and Stw 329, the smallest fossil considered here, are ordinated away from the other specimens in opposite directions along the same axis. Therefore, Stw 505's pattern of morphological variation from the remainder of the fossil sample is viewed here as likely to be allometric in nature. The analyses in the following sections primarily focus on these two specimens, with secondary attention to Sts 25, Sts 71, Stw 13, Stw 329, Sts 26,

Stw 266, Stw 580, and Stw 187. Morphometric results clearly can provide useful insights into patterns of fossil variation. Because an allometric shape correction could not be derived from either species of *Pan*, however (Chapter 3), morphometric distinctiveness by itself is not viewed here as indicating species-level variation. The pattern of variation in traditional morphological characters was therefore used to confirm or refute the taxonomic implications of morphometric distinctiveness.

Ordinations of morphological observations

Binary morphological observations were recorded as described in the Materials and Methods chapter, and the data are presented in Appendix 3. For analysis of the morphological data, Simple Matching similarities (Cheetham and Hazel, 1969) between specimens were obtained, and ordinations were attempted in order to visualize their relationships, as described in the Materials and Methods chapter. The Simple Matching similarities between pairs of specimens, numbers of observable characteristics on which the scores are based, and total number of observable characteristics on each specimen are shown in Appendix 4. The baseline sets of fossils for which the similarities were ordinated were the same as for the morphological data. The core group of specimens was MLD 37/38, Sts 5, Sts 19, Stw 505, and Sts 25. Ordination group 1 consisted of these plus TM 1511 and Sts 71; group 2 included the core group and Stw 329, and group 3 added Stw 13.

Ordinations of three main groups of specimens

Attempts to use the eigenvalue-based PCO technique for ordination of the binary data, as for the distances derived from the morphometric data, were unsuccessful. Almost all PCO runs returned a first eigenvalue equal to zero, and subsequent eigenvalues were negative. Only for one fossil grouping was any positive eigenvalue obtained, and this single value encompassed only a tiny percentage of the total variation within the matrix, as shown in Table 4.4. The PCO technique, therefore, was not suitable for production of a visually meaningful ordination of the interspecimen distances (Cailliez and Pagès, 1976).

Fossil set		**First eigenvalue, as percentage of total**
Core group of 5 specimens	alone	0
"	+ TM 1511 and Sts 71 (Group 1)	0.5% (c=6)[23] to 4.6% (c=1)
"	+ Stw 329 (Group 2)	0
"	+ Stw 13 (Group 3)	0
"	+ TM 1511	0
"	+ Sts 71	0

Table 4.4. Results of PCO analysis of Simple Matching distance matrices for morphological data.

Similarly, all attempts to ordinate these matrices with NMDS in *PAST* were unsuccessful. This is likely due to the fact that one of *PAST*'s initial eleven configurations for NMDS runs is the PCO result (Hammer et al., 2001), which involved many large negative eigenvalues. This particular result cannot be interpreted in three real

[23] See "Principal Coordinates Analysis" section above for discussion of the *c* coefficient.

dimensions as would be necessary for the construction of a 3D configuration; this situation is likely responsible for the failure to ordinate the specimens.

The *NTSYS* package (Applied Biostatistics Inc., 1998) can be directed to begin the NMDS iterations with a random configuration of data points, which eliminates the problems caused by imaginary eigenvectors. As with the ordinations of morphological data, these ordinations were repeated 100 times in order to minimize the importance of local-minima traps (discussed in Chapter 2). The results of these ordinations are displayed in Figures 4.10-4.12 below. Note that the multiple-runs approach to NMDS is critical, especially in *NTSYS*; the lowest-stress configurations were obtained in the 79th, 26th, and 48th runs for these three matrices respectively.

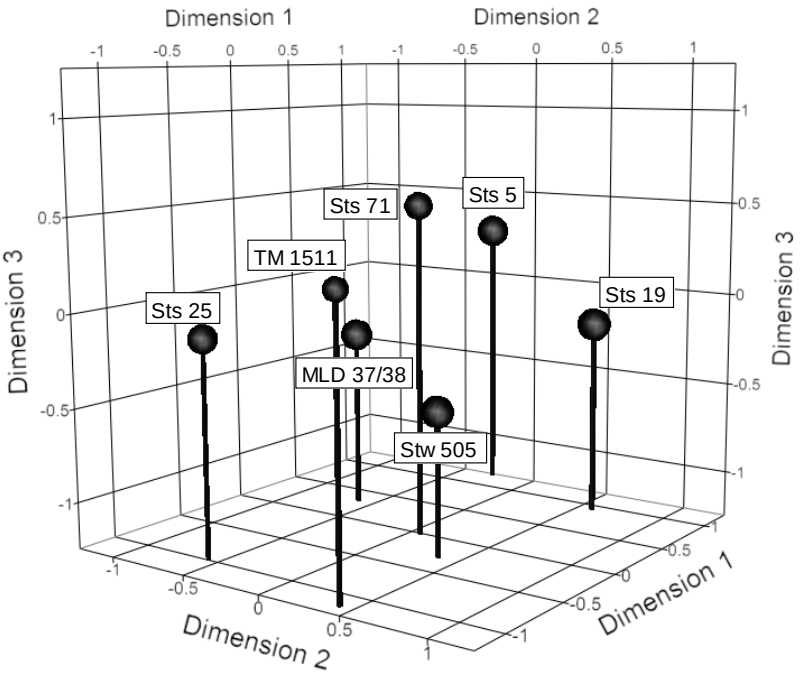

Figure 4.10. NMDS ordination of distances from morphological data for first set of specimens; stress = 0.2255.

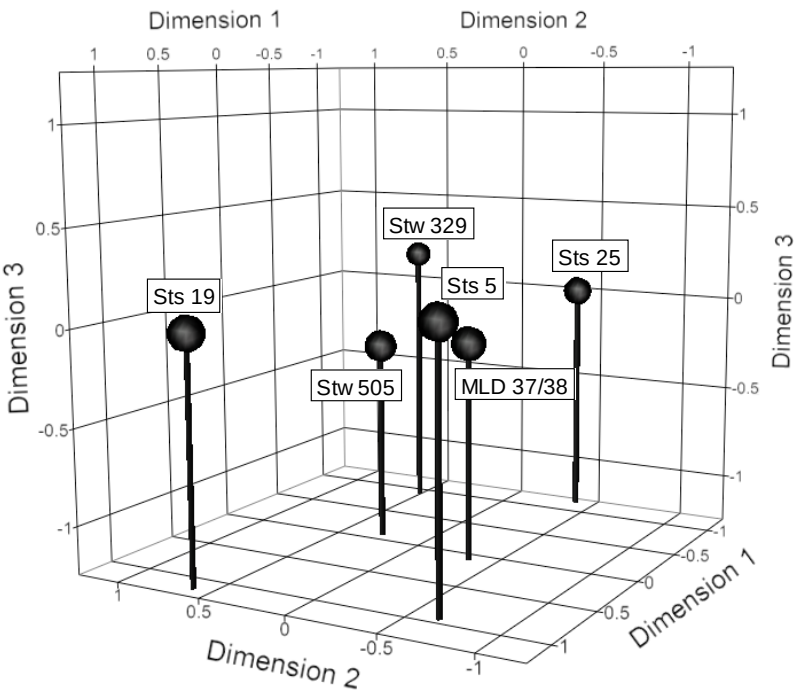

Figure 4.11. NMDS ordination of distances from morphological data for second set of specimens; stress = 0.0046.

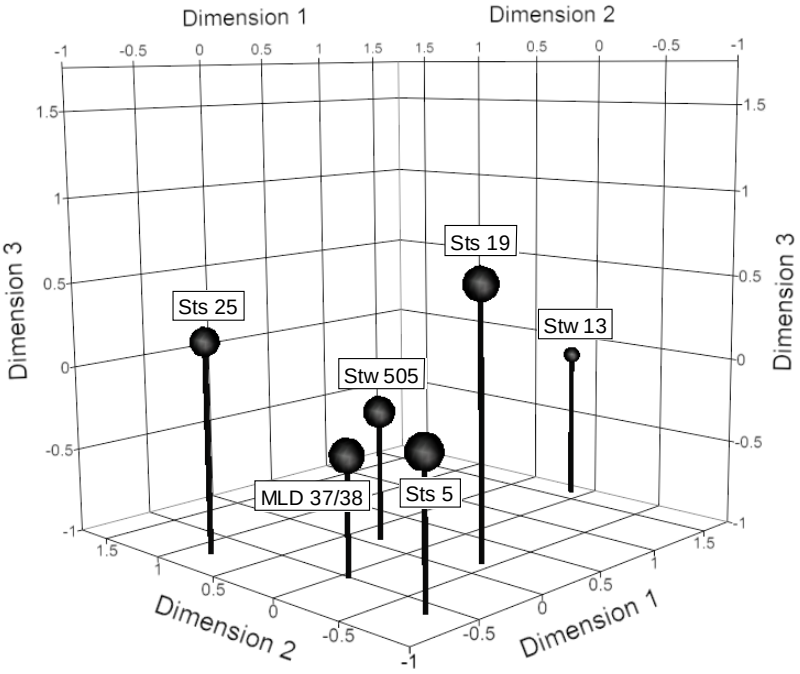

Figure 4.12. NMDS ordination of distances from morphological data for third set of specimens; stress = 0.0149.

As discussed in "Summary of Ordination Results" above, the specimens identified with morphometric methods as most clearly distinct from *A. africanus* were Stw 505 and Sts 19, and to a lesser extent Stw 13 and Sts 71. In Figures 4.10-4.12, we see that Stw 505 is consistently ordinated within the main group of specimens. As seen in Appendix 4, its similarities with the other specimens are generally moderate to high. This result is at odds with the result from the ordinations of morphometric data (Figures 4.2-4.7), in which Stw 505 was markedly distinct from the other specimens. This finding obviates the question above about whether Stw 13 tends to associate more closely with Stw 505 or the main group of specimens. It is interesting that Stw 13 is portrayed at the margin of the distribution in Figure 4.12, but its small number of observable characteristics make it difficult to draw firm conclusions about its taxonomy.

As noted above, ordinations of the morphometric data placed Stw 505 in distinction from the other specimens, but ordinations of the morphological observations do not. Further, given its unusual size, it is expected to have a different shape than the others. By the criteria described in Chapter 2, this indicates that Stw 505 is not considered likely to belong to a second species. The parsimonious inference is that Stw 505's morphometrics differ from the others due to allometry.

In contrast, Sts 19, which was somewhat distinct from the other specimens in the morphometric data ordinations along axes orthogonal to the one separating Stw 505 from the rest, remains distinct from the others here. Its similarities with the other specimens are generally moderate to low, despite sharing a large number of observable traits with them (Appendix 4). It is therefore distinguishable from the specimens traditionally

attributed to *A. africanus* both morphometrically and with respect to traditional morphological observations.

It is noted above ("Key specimens identified with morphometric techniques") that the ordinations of morphometric distance matrices associated Sts 71 with Sts 19. In the ordination of morphological distances (Figure 4.10), Sts 71 aligns more closely with the main group of specimens, although the ordination depicted in that figure has a very high stress value (0.2255). The high degree of stress makes the ordination presented in Figure 4.10 difficult to interpret. In Appendix 4, it can be seen that Sts 71 and Sts 19 share a moderately strong similarity over twelve common observable characteristics, but Sts 71 is typically more similar to the other specimens than is Sts 19. Figure 4.1 also shows that the Procrustes distance between these specimens is at approximately the 35th percentile of *Pan* pairwise distances. Sts 19 is included in many fossil pairs with high Procrustes distances, but Sts 71 generally has moderate distances from the other specimens, and is involved in no distances that meet the 95th percentile of *Pan* distances.

Sts 25 also stands in distinction from the other specimens in all three of these ordinations, just as it did in the ordinations of morphometric data. As with Sts 19, the consistency of this result calls its taxonomy into question.

Ordination of sets involving additional specimens

Finally, it was also noted in the section referenced above that several other specimens seemed to merit additional attention; namely, Sts 26, Stw 187, Stw 266, Stw 329, and Stw 580. Two additional interspecimen simple-matching similarity matrices were constructed for their ordination. Group 4 consists of the "core group" referenced

above as well as Stw 266, Stw 329, Sts 71, Stw 98, and MLD 31. The latter two specimens, though not included in the list above, were included here because they happened to have observable characteristics in common with all of the others in this group. These specimens are ordinated in Figure 4.13. For Group 5, not all members of the core group could be included because they did not all have characteristics in common with the remaining specimens (most notably, Sts 25). Group 5 therefore consisted of MLD 37/38, Sts 5, Sts 19, Sts 26, Stw 187, and Stw 580. These specimens' similarities are ordinated in Figure 4.14. As with all other efforts to ordinate matrices based on Simple Matching similarities, PCO was not suitable, as the first axes for these two sets described less than 12% of the variation regardless of *c* value, the second axes were approximately zero, and the third axes were zero or negative. NMDS was therefore the only ordination technique available.

The stress value associated with the ordination in Figure 4.13, 0.3211 (obtained on the 92nd of 100 iterations), is quite high. Obtaining a stress value this high with even 10 data points for 3 dimensions would be extremely unusual with random data (Klahr, 1969), indicating the presence of some internal contradictions within the similarity matrix. This is almost certainly the effect of missing data, and one must interpret this result with caution. Nonetheless, both Sts 19 and Sts 25 remain at the margins of the configuration, although less clearly so than in the previous figures. We have also already seen that Sts 71 and Stw 329 have been ordinated among the main distribution of specimens; that result is repeated here. MLD 31 similarly takes an unremarkable location in the configuration. The most noticeable result is the extent to which Stw 98 is distinguished from the other specimens, and to a lesser extent, the same is true of Stw

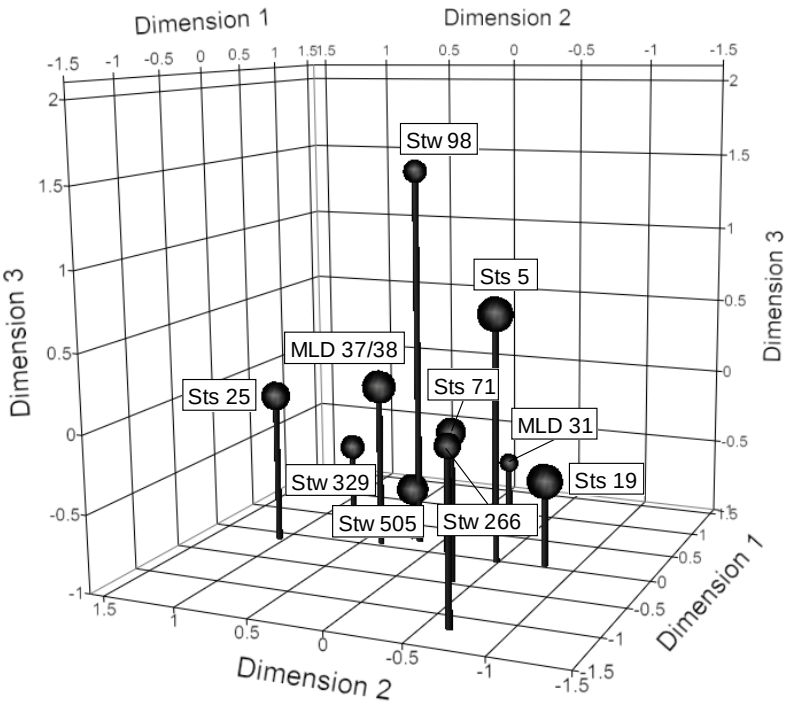

Figure 4.13. NMDS ordination of distances from morphological data of Group 4 specimens; stress = 0.3211.

266. These specimens unfortunately had too few morphometric landmarks in common with other specimens to allow their consideration in the morphometric ordinations above.

The problems with this ordination can be mitigated by referring to the original data on which it was based. The data presented in Appendix 4 show that Stw 98 shares moderate to high similarities with the majority of the other specimens, excluding Stw 266 (with which it shares only 2 observable traits), Stw 505, and Sts 19. The low similarity with Stw 505 appears to be the primary factor responsible for its apparent distinction. The specimens with which it shares the most observable characteristics are MLD 37/38, Sts 5, and Sts 19; with six, seven, and seven common observable landmarks respectively.

It is more similar to MLD 37/38 and Sts 5 than it is to Sts 19. Therefore, these results indicate that it would likely be inappropriate to separate Stw 98 from *A. africanus*.

Similarly, Stw 266 also appears to be somewhat distinct in this ordination. This result is likely due to its perfect dissimilarity with MLD 31 and Stw 98 on one and two observable characteristics, respectively; plus its perfect similarity with Sts 71, also based on one observable trait. As with Stw 98, this apparent distinctiveness is difficult to evaluate due to the small sample of observable characteristics, but the parsimonious explanation is that it is mainly due to the inflated effect of a few similarity scores based on little underlying data, and therefore not supportive of a split.

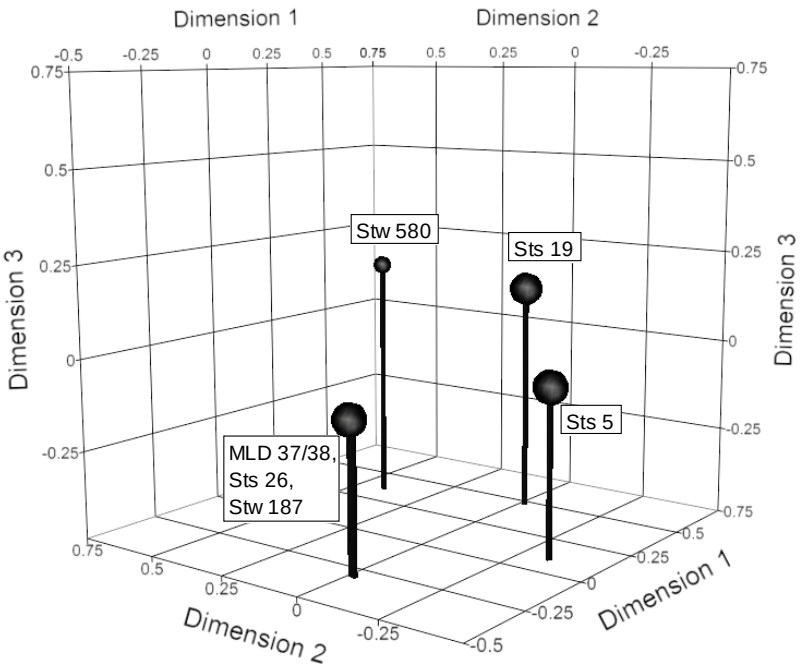

Figure 4.14. NMDS ordination of distances from morphological data of Group 5 specimens; stress = 0.142.

In Figure 4.14, it can be seen that Sts 19 remains distinct from Sts 5 and MLD 37/38. The stress in this figure is 0.142, which was obtained on all 100 ordination

iterations. Sts 26 and Stw 187 do not differ in the four traits on which they could be compared, and each of them also exhibits the same state on two common traits with MLD 37/38, so their "pins" are coincident on this figure. The main finding of interest in this ordination, however, is the extent to which Stw 580 is clearly distinct from most of these specimens, and somewhat allied with Sts 19. This corroborates the results shown in Figures 4.1 and 4.8, which indicate that Stw 580 is highly dissimilar to the other specimens, although it is based on a low number of observable morphological traits (Stw 580 only has four).

Unlike the situation with Figure 4.13, the acceptable stress level with this ordination (0.142, obtained on all 100 NMDS runs in *PAST*) indicates that the visual representation does not differ substantially from the underlying "true" interspecimen distances. These comparisons are complicated by the fact that the similarity indices are often based on very small numbers of comparable characteristics (see Appendix 4). Stw 580 is more similar to Sts 19 than it is to the others, and Sts 19 in turn is also distinct from the other specimens. The three tightly clustered specimens are all as similar to one another as possible, given the characteristics they have in common. To the extent that the fossils preserve informative morphology, these results are *consistent* with the hypothesis that Stw 580 is distinct from the remainder of the fossils, but with so few preserved characteristics, are not *dispositive*.

Specimens likely distinct from A. africanus

The results of these analyses indicate that the specimen most likely to be distinct from *A. africanus* is Sts 19 (contra Ahern, 1998). Both types of ordination of the morphometric data, the index of variability, and the ordination of the morphological observations consistently show that Sts 19 does not belong in the *A. africanus* taxon. The morphometric results had also indicated that Stw 505 bore substantial shape distinctions from the other specimens, but this finding was not borne out by the morphological observations, and appears to be the result of allometry. Since Stw 505 is the largest specimen considered here, the parsimonious inference is that the shape distinction is due to the effects of allometry.

Although its morphometric shape is only moderately distinct from the others, the morphological characteristics of Sts 25 are also consistent with it belonging to a taxon other than *A. africanus*. Finally, Stw 580 preserves a relatively small fragment of occipital, and therefore only a few recordable morphometric landmarks or morphological traits. The morphology that is preserved, however, consistently indicates that it is unlike the other specimens considered here. The percentile scores for its Procrustes distances from other specimens are very high, as are the average index of variability for tests in which it was included, and its particular pattern of morphological characters resulted in its ordination far from the other specimens. The implications of these distinctions are discussed further in the next chapter.

Chapter 5. Conclusion

The variability within the fossils traditionally assigned to *A. africanus* is inconsistent with the hypothesis of a single species. Three specimens in particular stand out in both morphometric and morphological analyses: Sts 19, Sts 25, and Stw 580. As discussed below, since the traditional appreciation of *A. africanus* from Sterkfontein is that it is limited to Member 4 (Robinson, 1954) and possibly Member 2 (Clarke, 2008)[24] as opposed to Member 5 (Curnoe and Tobias, 2006), the uncertainty over the provenience of Sts 19 indicates that it alone cannot upend our taxonomy of *A. africanus*.

Proponents of polytaxy within *A. africanus* typically come to one of two conclusions: identifying a small number of "exceptional specimens" (Lockwood, 1997:299), or separating the sample into two groups with approximately equal numbers of specimens. Clarke (e.g. 1994; 1988) is the most vocal proponent of the latter approach. The results here are consistent with the hypothesis that a second taxon is represented in the *A. africanus* sample, but in small numbers. Clarke's taxonomy is not supported by these results. This can be seen in Chapter 4 in two ways. First, there are not consistent divisions between large groups of specimens with approximately equal numbers; rather, individual specimens ordinate outside a single main group. Second, Sts 71, whose craniofacial differences from Sts 5 have been discussed in detail by Clarke, is not consistently isolated in the ordinations as are Sts 19, Sts 25, and Stw 580.

Others have identified Stw 183 (Lockwood and Tobias, 2002; Moggi-Cecchi, 2001; Lockwood, 1997), Stw 252 (Spoor, 1993; cited in Lockwood, 1997), MLD 37/38 (Picq, 1990), Stw 151 (Moggi-Cecchi et al., 1998), and Stw 255 (Lockwood and Tobias,

[24] Clarke does not attribute the "Little Foot" skeleton to *A. africanus*, but rather to his proposed second species within the sample of specimens that others attribute to that taxon.

2002) as exceptional specimens. Because this project focused on anatomy not preserved by most of those fossils, the only one of them on which these results shed light is MLD 37/38, which is discussed in the "Makapansgat vs. Sterkfontein" section below.

This project included several fragmentary specimens of uncertain taxonomy, such as Stw 98, Stw 187, Stw 266, Stw 329, Stw 580, and MLD 31 (see especially Lockwood and Tobias, 2002). Of these, only Stw 580 consistently ordinates in a way that suggests isolation from *A. africanus*. A slight distinction between Stw 329 and the other fossils is apparent in the morphometric ordinations, but its very small size appears to account for this, given that the direction of its separation from the main group of specimens is diametrically opposed to that of Stw 505, the largest specimen considered here. The remaining fossils ordinate in association with the main group of *A. africanus*, indicating that they do not bear consistent differences from that taxon in either morphometrics or traditional morphological characters. Given this result and their provenience, the parsimonious inference is that they are likely to belong to *A. africanus*. The absence of sufficiently diagnostic morphology, however, renders this conclusion tentative. Lockwood (1997) has noted that the likelihood of a second taxon in Sterkfontein Member 4 contraindicates the automatic attribution of all Member 4 specimens, especially when fragmentary, to *A. africanus*. There is, however, no finding here that supports splitting them out of that species.

Lockwood and Tobias (2002) have argued against declaring new species based on cranial evidence alone (preferring that alpha taxonomy incorporate evidence from anatomical regions, such as the dentition, that were outside the scope of their report based on craniofacial anatomy). As this project involved the basicranium, this discussion will

follow their lead, and identify specimens which introduce an excess of variation to the *A. africanus* sample, as inferred from consistent morphometric and morphological results derived from the basicranium.

Effect of assumptions made in the analyses

As noted in Chapter 1, the genus *Pan* is viewed here as an appropriate modern reference for species-level variation in the *A. africanus* sample. The results of the variability index tests described in Chapter 4 indicate that the fossil basicrania in this sample are slightly more variable than are those of *Pan*. If it could be shown that *Pan* is an inappropriate referent, that finding would have to be reexamined. The identification of exceptional specimens, however, probably would not be affected to a great extent. The morphometric data were referred to the distribution of pairwise Procrustes distances within *Pan* in order to generate proxy distance measurements that could be generalized between tests. If we assume that a more variable group such as gorillas or orangutans would offer a better referent for that purpose, the *relative* differences between fossil specimens would be unchanged, even if their *absolute* scores (percentile rank of Procrustes distances) were reduced. This reduction would be most pronounced for specimens with very highly ranked Procrustes distances, but the amount of this reduction could not be known without an appropriate dataset. In any event, the overall pattern of relationships among the fossils would remain unchanged, because their relative distances would not be affected. The exceptional specimens would remain exceptional.

Missing data and the approach adopted here

As described in the Results chapter, the approach adopted here allows simultaneous consideration of more specimens via both morphometric and morphological techniques than would otherwise be possible because it uses indirect distance proxies. While this can introduce to the ordinations a degree of interpretive error associated with specimens that have few available observations, this problem can be reduced by using visualization software (such as *JMP*, SAS Institute, 2007) that allows the portrayal of a fourth variable to denote the number of observable landmarks or traits. This fourth variable can serve as a proxy for the degree of confidence in the specimen's ordinated position, and prevents the problem of missing data from being fatal to interpretation. The ordination figures in the Results chapter were created with this in mind.

Most of the ordinations would have been either impossible or limited to very small numbers of landmarks or observations because of the fossils' often non-overlapping preservation. Others have noted that taxonomic studies are best pursued with inclusion of intermediate specimens as opposed to extremes (e.g. Lockwood et al., 1996), and that observed variability is highly dependent on sample size (Trinkaus, 1990). This approach takes those considerations into account by maximizing specimen counts. Should the analyses have been limited to (for example) two well-preserved specimens, it would be difficult to evaluate the significance of any observed distinction without a distribution of apparently related specimens. As discussed in the Results chapter, comparison of the second ordination group of fossils would have been impossible without this approach that accommodates missing data.

The sequential use of morphometrics and traditional morphological observations strengthens the conclusions. Morphometrics offers a powerful and objective way to characterize and analyze shape differences between specimens (Slice, 2005b; Guy et al., 2003; Bookstein, 1991), but landmark sets can miss biologically-relevant information in the intervening spaces (Richtsmeier et al., 2002; Aiello et al., 2000), and the treatment of "size" remains somewhat controversial (Slice, 2005a; Bookstein, 1991; Lele and Richtsmeier, 1995). Traditional morphological traits, on the other hand, easily capture information about the shapes of individual features and the relationships between those features, surface orientations, and the expression of soft-tissue contact markers on bone. Unfortunately, as many traditional traits involve the conflation of continuous variation into discrete categories, they are highly prone to interobserver error (Collard and Wood, 2000; Ahern, 1998), making these observations difficult to replicate (e.g. Holloway, 1981; White and Falk, 1999) and therefore less than satisfying from a scientific standpoint.

Because these approaches are complementary, however, their sequential application mitigates many of these problems. Allometric variation that may be difficult to handle morphometrically for a small fossil sample is not necessarily reflected in the specimens' morphological characteristics, especially in the basicranium. Further, the objectivity offered by morphometrics reduces the seriousness of the interobserver error problem for morphological traits. Specimens that stand in distinction under both approaches thus merit extra attention as possibly belonging to different taxa. Conversely, the parsimonious inference for specimens that seem distinct morphometrically but not

under traditional observations is that they provide examples of intraspecific variation. In the case of Stw 505, that variation is likely due to allometry.

Comparison of results with others' arguments

Sts 19

The Sts 19 fossil has been a perennial topic of taxonomic disagreement, perhaps partly because its Member provenience at Sterkfontein is uncertain. Broom and Robinson (1950:27) note that it (their "Skull 8") was recovered from a quarrying dump that they estimated to have been in use 30 to 40 years previously. The location of this dump near the site where Sts 5 ("Skull 5") was recovered seems to have been a consideration in their conclusion that these specimens were conspecific (ibid., p. 70), especially given Dart's habit of viewing specimens from the same site as conspecific (Tobias, 2001). Dean and Wood (1982) note that Sts 19 is the most *Homo*-like of the fossils traditionally assigned to *A. africanus*, but assert that this is most likely due to intraspecific variation. Their finding echoed that of Broom and Robinson (1950:70), who noted that its differences from Sts 5 amounted to a "very considerable degree of variation in *Plesianthropus transvaalensis*." Schepers (1950), published in the same volume, did not disagree with Broom and Robinson's (op. cit.) taxonomic assessment but pointed out several similarities between the brain of Sts 19 and that of modern humans, as evidenced by its endocast. As discussed in Chapter 1, both Kimbel (1984; Kimbel and Rak, 1993) and Grine (Strait and Grine, 2004; Smith and Grine, 2008) have changed their minds about the attribution of this fossil, and in opposite directions.

In its general appearance, Sts 19 differs from the remainder of the sample examined here in several ways. Its exhibits a greater degree of cranial flexion, a wider glenoid fossa, a larger ectoglenoid process, and more coronally-oriented petrosals (especially when compared to Sts 5), for example. Its external auditory meatus is more vertically elliptical than those of the other specimens, and it has no eustachian process. There is an unusually wide (as much as 4 mm) gap between the inferior aspect of the petrosal and the basilar part of the occipital. The fossae for insertion of the longus capitis muscle are anteromedially-posterolaterally oriented, unlike the other specimens, where these fossae are elongated in a more anteroposterior direction.

Sherwood and colleagues (2002), for example, group Sts 19 with Sts 5 and MLD 37/38 because they all exhibit petrous crests, to the exclusion of Sts 25 and 71. The ordinations presented in the Results chapter, however, are not consistent with this proposal. Sts 25 and Sts 71 do not display consistent affinity for one another to the exclusion of the other specimens, and Sts 19 is typically either at the margin of the ordinations or well separated from the others. These data therefore do not support their argument.

Braga and Boesch (1997a) place Sts 19 in *A. africanus*, and Braga et al. (1998) point out that Sts 19 shares with other specimens of *A. africanus* and anatomically modern humans the derived pattern of having the foramen ovale located on the sphenoid bone instead of the sphenosquamosal suture. They do, however, note that others (Kimbel and Rak, 1993) find substantial differences between Sts 19 and *A. africanus*. Despite their finding, the results here (based partially on morphometric data involving the foramen ovale) do not link Sts 19 with *A. africanus* as a whole.

Others (Curnoe and Tobias, 2006) also consider Sts 19 to represent *A. africanus*, but point out some features that it has in common with Stw 53, which they assign to *Homo*. Given the consistent distinctions between Sts 19 and the other fossils considered here, that similarity bears further consideration. The Procrustes distance between these specimens, however, is at the 83.5th percentile of the all-*Pan* distribution of distances based on the same set of landmarks (Appendix 1). Further comparisons of these specimens with other early *Homo* specimens would help to support the inference that they belong to that genus.

Ahern (1998), addressing work by Kimbel and Rak (1993), disputes both the claim that Sts 19 is excessively different than the remainder of the *A. africanus* specimens[25], and some of the morphological observations that led to that conclusion. This project, however, includes the observations disputed by Ahern. Nonetheless, Sts 19 is consistently shown here to be distinct from the remainder of the fossils. It is the specimen involved in the largest number of pairwise interfossil Procrustes distances that exceed the 95th percentile of corresponding distances between *Pan* specimens, and ordination of its morphological observations consistently places it at the periphery of *A. africanus*, well away from an axis connecting Stw 505 to the main group. The parsimonious inference is therefore that this fossil is distinct from that species, and possibly was deposited in Member 5 of Sterkfontein, or perhaps the "Stw 53 Infill" (Kuman and Clarke, 2000), which then would be closer in age to Member 5 than it is to Member 4, as implied by the presence of cutmarks on the fossil (Pickering et al., 2000).

[25] Lockwood and Tobias (1999:678) note that the cast from which Ahern's (1998) data for Stw 505 were gathered depicted some matrix that obscured certain observations, and that further preparation of the fossil has indicated that at least one of his observations was in error. His data about *A. africanus* variability are therefore incomplete.

As pointed out by Lockwood (1997), however, the uncertainty over the provenience of this fossil prevents its use in support of a claim of Sterkfontein Member 4 polytaxy. This would matter because Member 5 is substantially younger than Member 4, and has been shown to contain stone tools while Member 4 does not (Partridge, 1978). Hominins from Member 5 are therefore unlikely to belong to the same species (or at least morphotype) as those from Member 4.

Finally, Sarmiento (1993) also examined basicranial characters in Sts 19 and notes the extensive differences between Sts 19 and Sts 5 (although, given the length of its cranial base and the shortness of Sts 19's, Sts 5 may offer the most extreme distinctions from Sts 19 of any Member 4 specimen). Interestingly, he also pointed out that it shares some features with the robust australopithecines of the Swartkrans sample. The Procrustes distance between an exemplar of the Swartkrans robust australopithecines (SK 47) and Sts 19 is in the 72nd percentile of the corresponding *Pan* distribution (Appendix 1). Sts 19 is therefore moderately distant from exemplars of both *Homo* and *A. robustus*. This makes the difficulty in obtaining a secure provenience and therefore a date for it particularly unfortunate, as it may be an early specimen from one of those taxa.

Sts 25

Sts 25 is a somewhat crushed cranial base. As discussed above, others have identified it as possibly distinct from *A. africanus*, or at least pointed out that it retains features different from the remainder of the sample. Kimbel, for example, has noted the differences between this fossil and most of the others traditionally associated with *A. africanus* (Kimbel and White, 1988), but later revised his view and limited his list of

possibly-distinct specimens to Sts 19 (Kimbel and Rak, 1993). Using an approach based on the standard error of the slope for linear regressions of homologous measurements between specimens (following Thackeray et al., 2000), Lee and Wolpoff (2005) also generated results indicating that Sts 25 is substantially distinct from both Sts 5 and Sts 71. This distinction is much greater than that between any pairs of Swartkrans or habiline specimens they studied, including KNM-ER 1813 and KNM-ER 1470, which appear to represent different taxa (e.g. Wood, 1999; Grine et al., 1996; but see Miller, 2000). Unfortunately, Schwartz and Tattersall (2003; 2005) omit Sts 25 from their discussions of the morphology and affinities of the Sterkfontein specimens.

Sherwood and colleagues (2002) noted that Sts 25 aligns with Sts 71 in lacking a tympanic crest, while Sts 5, Sts 19, and MLD 37/38 all retain this feature. The analyses conducted here do not support this linkage, however, because Sts 25 and Sts 71 are not consistently ordinated together. Similarly, the results of Dean and Wood (1982) portray Sts 25 as having more sagittally oriented petrosals like Sts 5, to the exclusion of Sts 19 and MLD 37/38, but a shortened sphenoid like MLD 37/38 and Sts 19, to the exclusion of Sts 5. As with Sts 71, the results here do not consistently support any of these groupings. This is not surprising, as Sts 25 has a morphological pattern unlike any of those specimens. Its *Pan*-distribution percentiles for Procrustes distances indicate that it is morphometrically most like the juvenile specimen Stw 329, but also similar to TM 1511, Sts 71, and MLD 37/38. Its morphological similarities are also highest with these specimens.

Despite these similarities, the consistency with which Sts 25 is ordinated at the periphery of the samples examined here, both morphometrically and morphologically,

indicates that Sts 25 may well belong to a taxon other than *A. africanus*. This set of similarities indicates two possibilities that are not necessarily mutually exclusive. One of these is that Sts 25 represents a precursor of the robust australopithecines, as Sts 71 and the Makapansgat specimens (including MLD 37/38) generally bear a stronger resemblance to that group than do the other *A. africanus* specimens, such as Sts 5 (Lockwood, 1997). Another possibility is that Sts 25 is one of the older Member 4 specimens, as the Makapansgat specimens are generally accepted as being older than the Sterkfontein sample (Vrba, 1982). This second possibility seems to be less likely because the Makapansgat material does not consistently ordinate in distinction of the Sterkfontein material, which would be the case if Sts 25 appeared distinct only due to geological age (see below). Its mixture of primitive (sagittally oriented petrosals, shallow and tubular tympanic plate) and derived (shortened sphenoid) traits, combined with its tiny eustachian process, make attribution to a species difficult, so the suggestion here is that its specific designation be dropped, so it would represent *Australopithecus sp. indet.*

Stw 580

Stw 580 is a fragment of the left side of an occipital. Because it preserves little basicranial anatomy and no facial or dental structures at all, it has received very little attention in the literature. The only discussion of it appears to be in Lockwood and Tobias's (2002) treatment of specimens from Sterkfontein Member 4. Those authors refer it to their Group B. This group designates specimens that have recognizable differences from the robust australopithecines, but otherwise cannot be assigned to a species because they lack sufficiently diagnostic morphology. They note that its occipital

condyles are shorter and broader than is often seen in early *Homo*, but generally similar to those seen in *A. africanus*; that its jugular process is superoinferiorly thicker than those of early *Homo* or robust australopithecines; and that it does not have a marginal sinus, unlike *A. afarensis*. Taung does have an enlarged O/M sinus, however (Falk, 1990), demonstrating that the specimens assigned to *A. africanus* are variable for this trait. Such variability indicates that this trait is not useful in isolation for identifying specimens of *A. africanus*.

In the present results, Stw 580's morphometrics indicate that it is very different from the other specimens, including both Sts 5 and Sts 19. Very few *Pan* pairs approach the Procrustes distance between Stw 580 and those specimens. Its morphological observations, meanwhile, are very similar to those of Sts 19, but the fragmentary nature of Stw 580 left only two observable traits in common with Sts 19. Stw 580 therefore may represent a new taxon at Sterkfontein, but its absence of the anatomy typically used for identification of new species indicates that the best assignment for it at present is *Australopithecus sp. indet.*, and like Sts 25, probably not *A. africanus*. Unfortunately, the degree of preservation precludes direct comparisons between Sts 25 and Stw 580.

Makapansgat vs. Sterkfontein

The idea that the Makapansgat specimens differ taxonomically from those at Sterkfontein is an old one, dating to the initial announcement of finds there (Dart, 1948a, b). Although this taxonomic approach largely fell out of favor after Robinson (1954) explicitly referred the Sterkfontein and Makapansgat material to the same *sub*species, the possibility of a site-based distinction remains under consideration. Lockwood (1997), for

example, noted this distinction and argued that while one could argue for a resurrection of *A. prometheus*, he simply supported the removal of a specific designation for the Makapansgat fossils (*Australopithecus sp. indet*.). These specimens are likely to be older than the Sterkfontein sample (Vrba, 1982), but exhibit some features in common with later robust australopithecines (e.g. Picq, 1990). This similarity has been noted by Schwartz and Tattersall (2005), whose "morphs" for Makapansgat fossils often include TM 1517, the holotype of *A. robustus*. Others, however, disagree, arguing that site plays relatively little role in *A. africanus* variability (Kimbel and White, 1988).

The results from this study can provide some insight to this question, but only two Makapansgat specimens retain sufficient preservation to bear on these analyses. The well-preserved MLD 37/38 is consistently ordinated near Sts 5 and the main group of fossils on both morphometric and morphological observations. This finding echoes that of Schwartz and Tattersall (2005:222-23), who assign MLD 37/38 and Sts 5 to the same "cranial morph" and later discuss their basicranial similarities. MLD 31 only preserves five of the observed morphological traits, and did not have sufficient numbers of morphometric landmarks in common to qualify for the ordinations of morphometric distances. Its morphological data are ordinated with the Group 4 set of specimens (see Figure 4.13). It is aligned between Sts 19 and the bulk of the specimens in that figure, but its low number of observable traits (indicated by the small size of the head for its "pin") and the high stress value for that ordination render interpretations based only on that figure rather problematic. Instead, interpretations of the results here with respect to the question of Makapansgat and Sterkfontein conspecificity must rely on the overall picture presented, and this in turn relies heavily on MLD 37/38. That specimen, as noted

above, consistently ordinates near Sts 5. To the extent that MLD 31 is informative, it fails to indicate affinity for MLD 37/38 to the exclusion of the other fossils. These results, then, do not support the hypothesis that the Makapansgat and Sterkfontein samples comprise different taxa.

Conclusions

The fossil known as Sts 19 displays morphometric and morphological distinctions from the remainder of the *A. africanus* sample, and these are consistent enough to warrant the inference that it belongs to a different taxon. As its provenience is uncertain, the parsimonious inference is that it was deposited in Sterkfontein Member 5, which dates to well after the last appearance of *A. africanus* otherwise. This specimen is thus not useful with respect to the question of whether *A. africanus*, as widely understood, comprises multiple species.

Other fossils of more secure provenience are more relevant to this issue. It does appear that a second taxon is represented among the Sterkfontein Member 4 fossils, represented by Sts 25 and possibly by Stw 580, although the latter specimen does not preserve sufficient anatomy to permit fruitful direct comparisons between them. Given its apparent rarity in the sample, it is possible that this second species was rare in terms of its overall demography. Alternatively, the variability inherent in the Plio-Pleistocene climate of South Africa (Strait et al., 2009) may have resulted in fluctuating ranges of *A. africanus* and another species so that the Sterkfontein site was only occasionally visited by the second species.

This project has shown that the problem of missing data in fossil taxonomic studies can be overcome by using morphometric data as a means of identifying specimens that introduce excessive shape variation into a sample. The next step is to examine the indicated fossils for morphological patterns that distinguish them from the remainder of the sample. A two-pronged approach such as this helps to minimize the effect of the drawbacks inherent in either method applied alone.

Appendices

Appendix 1. Morphometric tests on fossils

The following tables and text detail the results of the fossil morphometric variability measurements. Each set of specimens with six or more landmarks in common was analyzed together, as well as selected groups with four or five common landmarks. Groups with three or fewer common landmarks were not tested. Each basicranial specimen commonly attributed to *A. africanus* on which four or more landmarks were preserved was included in at least one test, as well as other Sterkfontein specimens of unknown or controversial affinity, and some exemplar South African specimens from other taxa (either *A. robustus* or early *Homo*).

As described for the "Taphonomic Error" section of the Materials and Methods chapter, the listing for each test includes the specimens considered, the landmarks present, and the membership of the *Pan* comparative samples. Also listed are the values of Relethford and Blangero's (1990) index of variability. As discussed in the Methods chapter, this index involves dividing the trace of the covariance matrix by the number of variables. The resulting value can be viewed as a proxy for "variability," and compared to that of another dataset. Finally, the Procrustes distances between specimens are shown, together with the percentiles of these distances in the overall *Pan* distributions of pairwise Procrustes distances. When a test involved three or more specimens, these results are presented in a table; results for groups of two specimens are presented in text.

Some well-preserved specimens are included in multiple analyses, so it was unavoidable that a given pair may be included in a number of different analyses. The interspecimen Procrustes distance result on the test that includes the greatest number of landmarks is highlighted below with boldface type.

Specimens Sts 5, MLD 37/38, and Sts 19 have approximately equal bilateral preservation of this region. Tests 17-19 (reported in the Taphonomic Error section of the Materials and Methods chapter) addressed the extent of bilateral symmetry present in each of them. Otherwise, only the left-side observations were used, except for Sts 5 in Tests 8, 12, 15, 23, and 24.

Test 1.
Specimens considered: Sts 5, MLD 37/38, Stw 53, SK 47
Landmarks: 4, 11, 12, 14, 15, 28, 30, 31
Pan sample[26]: Chimps: 14 Western, 16 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index values: all *Pan*: .000622, fossils: .000714
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	Stw 53	SK 47	MLD 37/38	Sts 5
SK 47	**.1929** **73.7** [27][28]			
MLD 37/38	**.1989** **77.4**	**.2270** **89.8**		
Sts 5	**.1757** **60.9**	**.2020** **79.0**	.1491 37.8	
Sts 19	**.2108** **83.5**	**.1906** **72.1**	.1069 7.7	.1666 53.5

Test 2.
Specimens considered: Sts 5, Sts 19
Landmarks: 1-4, 6-16, 18, 20, 21, 26-31
Pan sample: Chimps: 12 Western, 14 Central, 14 Eastern; Bonobos: 40
Relethford and Blangero (1990) index: all *Pan*: .000153, fossils: .000156
Procrustes distance between specimens, and percentile of this distance in all-*Pan* distribution of pairwise distances: **.1497, 57.7**

Test 3.
Specimens considered: Sts 5, MLD 37/38
Landmarks: 1-4, 9-16, 18, 20-24, 26-31
Pan sample: Chimps: 14 Western, 14 Central, 14 Eastern; Bonobos: 42
Relethford and Blangero (1990) index: all *Pan*: .000179, fossils: .000111
Procrustes distance between specimens, and percentile of this distance in all-*Pan* distribution of pairwise distances: **.1263, 12.3**

[26] Unless otherwise specified, each species and subspecies sample includes equal numbers of specimens by sex.
[27] Each cell reports the appropriate Procrustes distance, and the percentile rank of that distance in the test's all-*Pan* distribution, in that order.
[28] Boldface type for Procrustes distances and percentile ranks indicates a result based on the largest number of landmarks available for a given specimen pair.

Test 4.
Specimens considered: MLD 37/38, Stw 329, Sts 5
Landmarks: 4, 11, 12, 14, 15, 21, 22, 24, 28, 30, 31
Pan sample: Chimps: 14 Western, 14 Central, 16 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000570, fossils: .000567
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 329
Stw 329	**.2272** **79.7**	
Sts 5	.1497 23.7	**.1968** **61.6**

Test 5.
Specimens considered: MLD 37/38, Stw 98, Stw 329, Sts 5
Landmarks: 12, 14, 15, 22, 24, 28
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .001154, fossils: .001072
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 98	Stw 329
Stw 98	**.1916** **61.9**		
Stw 329	.2799 85.6	**.1337** **29.6**	
Sts 5	.1219 21.9	**.1663** **49.2**	.2388 76.5

Test 6.
Specimens considered: MLD 37/38, Stw 505, Sts 19
Landmarks: 9-11, 13, 14, 18-23, 26-28, 30
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000333, fossils: .000469
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 505
Stw 505	**.2147** **91.5**	
Sts 19	**.1847** **70.7**	**.2162** **92.2**

Test 7.
Specimens considered: MLD 37/38, Stw 505, Sts 5, Sts 19
Landmarks: 9-11, 13, 14, 18, 20-23, 26-28, 30
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000346, fossils: .000490
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 505	Sts 5
Stw 505	.2138 92.3		
Sts 5	.1630 46.2	**.2364** **97.4**	
Sts 19	.1794 67.4	.2228 95.2	.1942 81.6

Test 8.
Specimens considered: MLD 37/38, Sts 71, TM 1511, Sts 25, Sts 19, Sts 5 (right side)
Landmarks: 9, 17, 18, 19, 20, 26
Pan sample: Chimps: 4 female and 3 male Western, 8 Central, 3 female and 4 male Eastern; Bonobos: 22
Relethford and Blangero (1990) index: all *Pan*: .001008, fossils: .000685
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Sts 71	TM 1511	Sts 25	Sts 19
Sts 71	**.1126** **6.8**				
TM 1511	**.1687** **41.5**	**.2043** **68.6**			
Sts 25	.1225 10.7	**.1678** **40.8**	**.1269** **13.3**		
Sts 19	.1378 19.1	**.1389** **20.0**	**.1666** **39.9**	.1780 48.7	
Sts 5 right	.1557 31.2	**.2035** **67.5**	**.1536** **29.1**	.1697 42.6	.1126 6.8

Test 9.
Specimens considered: Stw 580, Stw 187, Sts 5, Sts 19
Landmarks: 6-8, 13, 16, 29
Pan sample: Chimps: 14 Western, 14 Central, 14 Eastern; Bonobos: 42
Relethford and Blangero (1990) index: all *Pan*: .001091, fossils: .002837
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	Stw 580	Stw 187	Sts 5
Stw 187	**.2921** **96.7**		
Sts 5	**.40772** **99.9+**	**.2953** **97**	
Sts 19	**.3601** **99.8**	**.3329** **99.2**	.1980 58.7

Test 10.
Specimens considered: Sts 26, Sts 5, Sts 19
Landmarks: 2, 3, 6-8, 13, 29
Pan sample: Chimps: 6 female and 7 male Western, 14 Central, 7 female and 6 male Eastern; Bonobos: 40
Relethford and Blangero (1990) index: all *Pan*: .000531, fossils: .000837
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	Sts 26	Sts 5
Sts 5	**.2032** **93.6**	
Sts 19	**.2248** **97.1**	.1185 22.7

Test 11.
Specimens considered: MLD 37/38, Sts 19, Sts 25
Landmarks: 4, 9-12, 18-22, 26-28
Pan sample: Chimps: 14 Western, 14 Central, 16 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000403, fossils: .000333
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Sts 19
Sts 19	.1742 51.3	
Sts 25	**.1452** **19.3**	**.1633** **39.4**

Test 12.
Specimens considered: MLD 37/38, Sts 5 (right side), Sts 19, Sts 25
Landmarks: 4, 9, 10, 12, 18-22, 26-28
Pan sample: Chimps: 14 Western, 16 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000403, fossils: .000417
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Sts 5 right	Sts 19
Sts 5 right	.1404 22.1		
Sts 19	.1696 56.0	.2111 89.5	
Sts 25	.1398 21.5	**.2016** **83.8**	.1649 50.9

Test 13.
Specimens considered: MLD 37/38, Sts 5 (left side), Sts 19, Sts 25
Landmarks: 4, 9-12, 18, 20-22, 26-28
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000453, fossils: .000345
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Sts 5 left	Sts 19
Sts 5 left	.1391 14.6		
Sts 19	.1662 41.6	.1706 46.3	
Sts 25	.1493 22	**.1651** **40.2**	.1540 26.6

Test 14.
Specimens considered: MLD 31, MLD 37/38, Stw 329, Sts 5
Landmarks: 15, 22, 24, 31
Pan sample: Chimps: 14 Western, 14 Central, 16 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .001976, fossils: .001873
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 31	MLD 37/38	Stw 329
MLD 37/38	**.2043** **65.6**		
Stw 329	**.1170** **35.7**	.2791 79.9	
Sts 5	**.2020** **65.1**	.0970 25.7	.2958 81.9

Test 15.
Specimens considered: Stw 13, Stw 505, Sts 5 (right side), Sts 19, Sts 25, TM 1517, MLD 37/38
Landmarks: 10, 18-20, 27
Pan sample: Chimps: 14 Western, 16 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .001147, fossils: .000747
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 13	Stw 505	Sts 5 (right)	Sts 25	Sts 19
Stw 13	**.1465** **35.2**					
Stw 505	.1700 50.9	**.0823** **3.3**				
Sts 5 right	.1379 29	**.1163** **16.5**	.1564 42			
Sts 25	.1421 32.2	**.1868** **61.7**	.2063 72.4	.1785 56.7		
Sts 19	.1481 36.6	**.0558** **0.2**	.1060 10.9	.1189 17.5	.1967 67.2	
TM 1517	**.1270** **21.7**	**.1476** **35.8**	**.1807** **57.8**	**.1341** **26.9**	**.1841** **60.6**	**.1258** **21.2**

Test 16.
Specimens considered: MLD 37/38, Stw 266, Sts 5, Sts 19
Landmarks: 11, 12, 21, 23, 28
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000773, fossils: .001396
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 266	Sts 5
Stw 266	**.1885** **82.8**		
Sts 5	.2003 87.6	**.2796** **99.0**	
Sts 19	.0814 6.3	**.2339** **95.1**	.1915 84.0

Tests 17-19.
These tests address the presence of bilateral symmetry in some well-preserved specimens, and are reported in the "Taphonomic Error" section of the Materials and Methods chapter.

Test 20.
Specimens considered: MLD 37/38, Stw 329, Stw 505, Sts 5, Sts 19
Landmarks: 11, 14, 21, 22, 28, 30
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000611, fossils: .000968
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 329	Stw 505	Sts 5
Stw 329	.1450 56.3			
Stw 505	.1880 87.7	**.2804** **99.6**		
Sts 5	.1258 35.7	.1406 51.8	.2343 97.7	
Sts 19	.1185 28.8	.2165 95.3	.1356 46.5	.2062 93.4

Test 21.
Specimens considered: MLD 37/38, Stw 329, Sts 5, Sts 19
Landmarks: 4, 11, 12, 14, 15, 21, 22, 28, 30, 31
Pan sample: Chimps: 14 Western, 16 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000417, fossils: .000453
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 329	Sts 5
Stw 329	.1598 57.7		
Sts 5	.1228 18.4	.1493 46.8	
Sts 19	.1442 41.2	**.2129** **94.2**	.18583 80.9

Test 22.
Specimens considered: MLD 37/38, Stw 329, Sts 5, Sts 19, Sts 25
Landmarks: 4, 11, 12, 21, 22, 28
Pan sample: Chimps: 14 Western, 14 Central, 16 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000496, fossils: .000492
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 329	Sts 5	Sts 19
Stw 329	.1383 63.5			
Sts 5	.10968 30.2	.1272 51.6		
Sts 19	.1155 37.3	.1577 79.7	.1490 72.7	
Sts 25	.1314 55.9	**.1010** **22.0**	.1568 79.0	.1328 57.0

Test 23.
Specimens considered: MLD 37/38, Stw 505, Sts 5 (right side), Sts 71, Sts 25
Landmarks: 9, 18-20, 22, 26
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000445, fossils: .000726
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 505	Sts 5	Sts 71
Stw 505	.1327 64.8			
Sts 5	.1446 77.3	.1619 89.4		
Sts 71	**.1313** **63.6**	**.1666** **91.5**	**.2526** **99.9**	
Sts 25	.1076 32.5	**.1507** **82.6**	.2094 99.1	**.1019** **25.6**

Test 24.
Specimens considered: MLD 37/38, Stw 505, Sts 5 (right side), Sts 19, Sts 25, Sts 71, TM 1511, TM 1517
Landmarks: 9, 18-20, 26
Pan sample: Chimps: 16 Western, 14 Central, 14 Eastern; Bonobos: 44
Relethford and Blangero (1990) index: all *Pan*: .000964, fossils: .000876
Procrustes distances between specimens, and percentiles of these distances in all-*Pan* distribution of pairwise distances:

	MLD 37/38	Stw 505	Sts 5	Sts 19	Sts 25	Sts 71	TM 1511
Stw 505	.1789 63.2						
Sts 5	.1580 48.1	.1355 32.0					
Sts 19	.1445 38.6	.1274 26.5	.1125 16.7				
Sts 25	.1174 19.8	.1912 71.4	.1967 74.8	.1724 58.5			
Sts 71	.0887 5.2	.2249 88.2	.2177 85.4	.1630 52.0	.1235 24.0		
TM 1511	.1599 49.6	**.1799** **63.6**	.1714 57.8	.1387 33.9	.1506 43.4	.1720 57.8	
TM 1517	.1710 57.8	.1885 70.0	.1076 13.4	.1069 13.0	.1833 66.1	.1946 73.6	.1617 51.2

Appendix 2. PCA results for **Pan** *shape variation*

This table shows the eigenvalues and associated variance percentages for the principal components analysis (PCA) performed on the *Pan* dataset.

PC	Eigenvalue	Percent of variance	Cumulative percentage
1	25.7118	16.695	16.695
2	13.1628	8.5469	25.2419
3	9.32881	6.0574	31.2993
4	8.30637	5.3935	36.6928
5	7.66487	4.977	41.6698
6	7.0132	4.5538	46.2236
7	5.93791	3.8556	50.0792
8	5.5617	3.6113	53.6905
9	4.94593	3.2115	56.902
10	4.40779	2.8621	59.7641
11	3.96958	2.5775	62.3416
12	3.73475	2.4251	64.7667
13	3.69012	2.3961	67.1628
14	3.3314	2.1632	69.326
15	3.12494	2.0291	71.3551
16	2.70278	1.755	73.1101
17	2.562	1.6636	74.7737
18	2.41857	1.5704	76.3441
19	2.12509	1.3799	77.724
20	2.05327	1.3332	79.0572
21	1.96235	1.2742	80.3314
22	1.90134	1.2346	81.566
23	1.77333	1.1515	82.7175
24	1.68808	1.0961	83.8136
25	1.6217	1.053	84.8666
26	1.41808	0.92079	85.78739
27	1.36987	0.88949	86.67688
28	1.32368	0.85949	87.53637
29	1.29254	0.83927	88.37564
30	1.17601	0.76361	89.13925
31	1.15826	0.75209	89.89134
32	1.05637	0.68592	90.57726
33	1.03102	0.66946	91.24672
34	0.87124	0.56571	91.81243
35	0.845948	0.54929	92.36172
36	0.798537	0.51851	92.88023
37	0.739937	0.48046	93.36069
38	0.699597	0.45426	93.81495

Continued next page

PC	Eigenvalue	Percent of variance	Cumulative percentage
39	0.657881	0.42718	94.24213
40	0.656884	0.42653	94.66866
41	0.622161	0.40398	95.07264
42	0.570852	0.37067	95.44331
43	0.532031	0.34546	95.78877
44	0.492241	0.31962	96.10839
45	0.461432	0.29962	96.40801
46	0.451964	0.29347	96.70148
47	0.420273	0.27289	96.97437
48	0.381621	0.24779	97.22216
49	0.334141	0.21696	97.43912
50	0.314549	0.20424	97.64336
51	0.304891	0.19797	97.84133
52	0.295723	0.19202	98.03335
53	0.273662	0.17769	98.21104
54	0.25151	0.16331	98.37435
55	0.242557	0.1575	98.53185
56	0.216792	0.14077	98.67262
57	0.210821	0.13689	98.80951
58	0.194589	0.12635	98.93586
59	0.174039	0.11301	99.04887
60	0.148891	0.096678	99.14555
61	0.140884	0.091479	99.23703
62	0.13298	0.086347	99.32337
63	0.121552	0.078926	99.4023
64	0.114238	0.074177	99.47648
65	0.105549	0.068535	99.54501
66	0.0905791	0.058815	99.60383
67	0.0859972	0.05584	99.65967
68	0.076628	0.049756	99.70942
69	0.0673459	0.043729	99.75315
70	0.0640585	0.041595	99.79475
71	0.0546763	0.035502	99.83025
72	0.0499297	0.03242	99.86267
73	0.045309	0.02942	99.89209
74	0.04238	0.027518	99.91961
75	0.0365305	0.02372	99.94333
76	0.0341086	0.022147	99.96547
77	0.0315136	0.020462	99.98594
78	0.0217012	0.014091	~100
79	2.32177E-12	1.5076E-12	~100
80	1.69924E-12	1.1034E-12	~100
81	1.38793E-12	9.0122E-13	~100
82	7.09852E-13	4.6092E-13	~100

Continued next page

PC	Eigenvalue	Percent of variance	Cumulative percentage
83	1.15744E-13	7.5155E-14	~100
84	7.98328E-14	5.1837E-14	100

Appendix 3. Morphological observations

Binary morphological data. See Chapter 2 for trait definitions. Blank cells indicate that a given trait is unobservable on a particular specimen.

Trait	MLD 37/38	MLD 31	TM 1511	Sts 5	Sts 19	Sts 25	Sts 26	Sts 71	Stw 13	Stw 98	Stw 187	Stw 266	Stw 329	Stw 505	Stw 580
1	1			1	0			0						0	
2	1		1	1	1	1		1	0					1	
3	1			1	1	0						0	1	0	
4	0			0	0	1		0				0	1	1	
5		1		1	1	0				1		0		1	
6	1			1	0	1				1			1		
7	1			1	1			1		1				0	
8	0	1		0	1					0			1		
9	1			0	1	0						0	0	1	
10	1	1		1	1	1		1		0			0	1	
11	0		0	1	1			0	0						
12	0			0	0	0		0						0	
13	1			1	1	0				0		1	1	1	
14	0			1	1	0						1	0	0	
15	0		1	1	1	0		1	1					1	
16	0			0	0			1						0	
17				1	1		0				0				1
18				1	0		1				1				0
19	0			0			0				0				1
20	1		0	1	0	0		0						1	
21	0	0		1	0					1			0	0	
22	1		0	1	1		1				1				
23	1	0												1	
24											1				0
25	0			1	0	0						1	0	0	
26	0				1			1						0	
27	0		0	0	1	0		0						0	

MLD 31	Stw 98	Stw 187	Stw 580	Stw 266	Sts 26	Stw 13	Stw 329	Sts 71	TM 1511	Sts 25	Stw 505	Sts 19	Sts 5	MLD 37/38	
0.5	0.5	1.0	0.0	0.5	1.0	0.333	0.6	0.583	0.5	0.615	0.722	0.571	0.714	**21**	**MLD 37/38**
0.5	0.714	0.75	0.333	0.571	0.75	0.333	0.4	0.636	0.5	0.429	0.529	0.609	**24**	21	**Sts 5**
1.0	0.286	0.333	1.0	0.571	0.333	0.333	0.5	0.75	0.5	0.357	0.611	**24**	23	21	**Sts 19**
0.75	0.2	na	na	0.429	na	0.5	0.625	0.546	0.75	0.615	**19**	18	18	18	**Stw 505**
0.5	0.5	na	na	0.286	na	0.0	0.625	0.714	0.75	**14**	13	14	13	13	**Sts 25**
na	na	0.0	na	na	0.0	0.667	na	1.0	**6**	4	4	6	6	6	**TM 1511**
1.0	0.5	na	na	1.0	na	0.667	0.0	**12**	5	7	11	12	11	12	**Sts 71**
0.667	0.4	na	na	0.167	na	na	**10**	2	0	8	8	10	10	10	**Stw 329**
na	na	na	na	na	na	**3**	0	3	3	2	2	3	3	3	**Stw 13**
na	na	1.0	0.0	na	**4**	0	0	0	1	0	0	3	4	2	**Sts 26**
0.0	0.0	na	na	**7**	0	0	6	1	0	7	7	7	7	6	**Stw 266**
na	na	0.0	**4**	0	3	0	0	0	0	0	0	2	3	1	**Stw 580**
na	na	**5**	4	0	4	0	0	0	1	0	0	3	4	2	**Stw 187**
0.25	**7**	0	0	2	0	0	5	2	0	4	5	7	7	6	**Stw 98**
5	4	0	0	1	0	0	3	1	0	2	4	4	4	4	**MLD 31**

Appendix 4. Compilation of data from morphological observations. Main diagonal (bold type): number of observable characteristics on a given specimen. At left: Simple Matching similarity scores (Cheetham and Hazel, 1969) between pairs. Right: Number of observable characteristics held in common by pairs (regardless of similarity). Undefined similarities (due to having zero common observable characteristics) are indicated by "na".

References

Ackermann RR. 2002. Patterns of covariation in the hominoid craniofacial skeleton: implications for paleoanthropological models. J Hum Evol 43:167-187.

Ahern JCM. 1998. Underestimating intraspecific variation: the problem with excluding Sts 19 from *Australopithecus africanus*. Am J Phys Anthropol 105:461-480.

Ahern JCM, Hawks JD, Lee S-H. 2005. Neandertal taxonomy reconsidered...again: a response to Harvati et al. (2004). J Hum Evol 48:647-652.

Aiello L, Dean C. 1990. An introduction to human evolutionary anatomy. London: Academic Press Limited.

Aiello LC, Collard M, Thackeray JF, Wood BA. 2000. Assessing exact randomization-based methods for determining the taxonomic significance of variability in the human fossil record. S Afr J Sci 96:179-183.

Albrecht GH. 1979. The study of biological versus statistical variation in multivariate morphometrics: the descriptive use of multiple regression analysis. Syst Zool 28:338-344.

Albrecht GH. 2000. Multivariate morphometric patterns of sexual dimorphism in primates. Am J Phys Anthropol [Suppl] 30:94.

Albrecht GH, Miller JMA. 1993. Geographic variation in primates: a review with implications for interpreting fossils. In: Kimbel WH, Martin LB, editors. Species, species concepts, and primate evolution. New York: Plenum Press. p 123-161.

Albrecht GH, Miller JMA. 1997. Variability profiles in fossil studies: the case of *Homo habilis*. Am J Phys Anthropol [Suppl] 24:65.

Applied Biostatistics Inc. 1998. NTSYS-PC numerical taxonomy and multivariate analysis program version 2.02i.

Asfaw B, White TD, Lovejoy CO, Latimer B, Simpson S, Suwa G. 1999. *Australopithecus garhi*: a new species of early hominid from Ethiopia. Science 284:629-635.

Benton MJ, Pearson PN. 2001. Speciation in the fossil record. Trends Ecol Evol 16:405-411.

Berger LR, Keyser AW, Tobias PV. 1993. Brief communication: Gladysvale: first early hominid site discovered in South Africa since 1948. Am J Phys Anthropol 92:107-111.

Berger LR, Lacruz R, de Ruiter DJ. 2002. Revised age estimates of *Australopithecus*-bearing deposits of Sterkfontein, South Africa. Am J Phys Anthropol 119:192-197.

Berger LR, Tobias PV. 1996. A chimpanzee-like tibia from Sterkfontein, South Africa and its implications for the interpretation of bipedalism in *Australopithecus africanus*. J Hum Evol 30:343-348.

Bookstein FL. 1991. Morphometric tools for landmark data: Geometry and biology. Cambridge: Cambridge University Press.

Braga J. 1998. Chimpanzee variation facilitates the interpretation of the incisive suture closure in South African Plio-Pleistocene hominids. Am J Phys Anthropol 105:121-135.

Braga J, Boesch C. 1997a. Further data about venous channels in South African Plio-Pleistocene hominids. J Hum Evol 33:423-447.

Braga J, Boesch C. 1997b. The “radiator” bias. A reply to Falk & Gage. J Hum Evol 33:503-506.

Braga J, Crubézy E, Elyaqtine M. 1998. The posterior border of the sphenoid greater wing and its phylogenetic usefulness in human evolution. Am J Phys Anthropol 107:387-399.

Braga J, Hublin JJ. 1998. What do carotid canals tell us about human brain evolution? Am J Phys Anthropol [Suppl] 26:112.

Broom R. 1936. A new fossil anthropoid skull from South Africa. Nature 138:486-488.

Broom R. 1938. Further evidence on the structure of the South African Pleistocene anthropoids. Nature 142:897-899.

Broom R, Robinson JT. 1950. Further evidence of the structure of the Sterkfontein ape-man, *Plesianthropus*. Pretoria: Transvaal Museum Memoir No. 4, Pt. 1.

Broom R, Schepers GWH. 1946. The South African fossil ape-men, the Australopithecinae. Pretoria: Transvaal Museum Memoir No. 2.

Brunet M, Beauvilain A, Coppens Y, Pilbeam D. 1995. The first australopithecine 2,500 kilometers west of the Rift Valley (Chad). Nature 378:273-274.

Bush AM, Powell MG, Arnold WS, Bert TM, Daley GM. 2002. Time-averaging, evolution, and morphologic variation. Paleobiology 28:9-25.

Cailliez F, Pagès J-P. 1976. Introduction à l’analyse des données. Paris: Société de Mathématiques Appliquées et de Sciences Humaines.

Calcagno JM, Cope DA, Lacy MG, Tobias PV. 1997. Is *A. africanus* the only hominid species in Sterkfontein Member 4? Am J Phys Anthropol [Suppl] 24:86-87.

Cameron DW. 2003. Early hominin speciation at the Plio/Pleistocene transition. HOMO - Journal of Comparative Human Biology 54:1-28.

Carlson KJ, Grine FE, Pearson OM. 2007. Robusticity and sexual dimorphism in the postcranium of modern hunter-gatherers from Australia. Am J Phys Anthropol 134:9-23.

Cheetham AH, Hazel JE. 1969. Binary (presence-absence) similarity coefficients. J Paleontol 43:1130-1136.

Clarke RJ. 1988. A new *Australopithecus* cranium from Sterkfontein and its bearing on the ancestry of *Paranthropus*. In: Grine FE, editor. Evolutionary history of the "robust" australopithecines. New York: Aldine de Gruyter. p 285-292.

Clarke RJ. 1994. Advances in understanding the craniofacial anatomy of South African early hominids. In: Corruccini RS, Ciochon RL, editors. Integrative pathways to the past: paleoanthropological advances in honor of F. Clark Howell. Englewood Cliffs, NJ: Prentice Hall. p 205-222.

Clarke RJ. 2008. Latest information on Sterkfontein's *Australopithecus* skeleton and a new look at *Australopithecus*. S Afr J Sci 104:443-449.

Clarke RJ, Partridge TC. 2002. On the unrealistic 'revised age estimates' for Sterkfontein. S Afr J Sci 98:415-419.

Cobb SN, O'Higgins P. 2004. Hominins do not share a common postnatal facial ontogenetic shape trajectory. J Exp Zool (Mol Dev Evol) 302B:302-321.

Collard M, Wood B. 2000. How reliable are human phylogenetic hypotheses? Proc Natl Acad Sci U S A 97:5003-5006.

Coolidge HJJ. 1933. *Pan paniscus*. Pigmy chimpanzee from south of the Congo river. Am J Phys Anthropol 18:1-57.

Cope DA, Lacy MG. 1992. Falsification of a single species hypothesis using the coefficient of variation: a simulation approach. Am J Phys Anthropol 89:359-378.

Corruccini RS. 1983. Principal components for allometric analysis. Am J Phys Anthropol 60:451-453.

Corruccini RS. 1987. Shape in morphometrics: comparative analyses. Am J Phys Anthropol 73:289-304.

Corruccini RS, McHenry HM. 1979. Morphological affinities of *Pan paniscus*. Science 204:1341-1343.

Curnoe D, Tobias PV. 2006. Description, new reconstruction, comparative anatomy, and classification of the Sterkfontein Stw 53 cranium, with discussions about the taxonomy of other southern African early *Homo* remains. J Hum Evol 50:36-77.

Dart RA. 1925. *Australopithecus africanus*: the man-ape of South Africa. Nature 115:195-199.

Dart RA. 1948a. The adolescent mandible of *Australopithecus prometheus.* Am J Phys Anthropol 6:391-412.

Dart RA. 1948b. The Makapansgat proto-human *Australopithecus prometheus*. Am J Phys Anthropol 6:259-284.

de Ruiter DJ, Steininger CM, Berger LR. 2006. A cranial base of *Australopithecus robustus* from the hanging remnant of Swartkrans, South Africa. Am J Phys Anthropol 130:435-444.

Dean MC. 1988. Growth processes in the cranial base of hominoids and their bearing on morphological similarities that exist in the cranial base of *Homo* and *Paranthropus*. In: Grine FE, editor. Evolutionary history of the "robust" australopithecines. New York: Aldine de Gruyter. p 107-112.

Dean MC, Wood BA. 1981. Metrical analysis of the basicranium of extant hominoids and Australopithecus. Am. J. Phys. Anthrop. 54:63-71.

Dean MC, Wood BA. 1982. Basicranial anatomy of Plio-Pleistocene hominids from East and South Africa. Am J Phys Anthropol 59:157-174.

Delson E. 1987. Evolution and palaeobiology of robust *Australopithecus*. Nature 327:654-655.

Delson E. 1988. Chronology of the South African australopith site units. In: Grine FE, editor. Evolutionary history of the "robust" australopithecines. New York: Aldine de Gruyter. p 317-324.

Digby PGN, Kempton RA. 1987. Multivariate analysis of ecological communities. London: Chapman and Hall.

Dolhinow P, Fuentes A (eds.) 1999. The nonhuman primates. Mountain View, CA: Mayfield Publishing Company.

Dryden IL, Mardia KV. 1998. Statistical shape analysis. Chichester, England: Wiley.

Eldredge N, Gould SJ. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, editor. Models in paleobiology. San Francisco: Freeman, Cooper and Company. p 82-115.

Falk D. 1986. Evolution of cranial blood drainage in hominids: enlarged occipital/marginal sinuses and emissary foramina. Am J Phys Anthropol 70:311-324.

Falk D. 1988. Enlarged occipital/marginal sinuses and emissary foramina: their significance in hominid evolution. In: Grine FE, editor. Evolutionary history of the "robust" australopithecines. New York: Aldine de Gruyter. p 85-96.

Falk D. 1990. Brain evolution in *Homo*: the "radiator" theory. Behav Brain Sci 13:333-381.

Falk D, Conroy GC. 1983. The cranial venous sinus system in *Australopithecus afarensis*. Nature 306:779-781.

Falk D, Gage TB. 1997. Flushing the radiator? A reply to Braga & Boesch. J Hum Evol 33:495-502.

Falk D, Gage TB. 1998. Radiators are cool: a response to Braga & Boesch's published paper and reply. J Hum Evol 35:307-312.

Falk D, Gage TB, Dudek B, Olson TR. 1995. Did more than one species of hominid coexist before 3·0 ma?: evidence from blood and teeth. J Hum Evol 29:591-600.

Fruth B, Hohmann G, McGrew WC. 1999. The *Pan* species. In: Dolhinow P, Fuentes A, editors. The nonhuman primates. Mountain View, CA: Mayfield Publishing Company. p 64-73.

Gapert R, Black S, Last J. 2009. Sex determination from the occipital condyle: Discriminant function analysis in an eighteenth and nineteenth century British sample. Am J Phys Anthropol 138:384-394.

Gauld SC. 1996. Allometric patterns of cranial bone thickness in fossil hominids. Am J Phys Anthropol 100:411-426.

Giles E. 1956. Cranial allometry in the great apes. Human Biol 28:43-58.

Gordon AD, Green DJ, Richmond BG. 2008. Strong postcranial size dimorphism in *Australopithecus afarensis*: results from two new resampling methods for multivariate data sets with missing data. Am J Phys Anthropol 135:311-328.

Gower JC, Legendre P. 1986. Metric and Euclidean properties of dissimilarity coefficients. Journal of Classification 3:5-48.

Grine FE, Jungers WL, Schultz J. 1996. Phenetic affinities among early *Homo* crania from East and South Africa. J Hum Evol 30:189-225.

Grine FE, Ungar PS, Teaford MF, El-Zaatari S. 2006. Molar microwear in *Praeanthropus afarensis*: Evidence for dietary stasis through time and under diverse paleoecological conditions. J Hum Evol 51:297-319.

Gunz P. 2005. Statistical & geometric reconstruction of hominid crania: reconstructing australopithecine ontogeny. Ph.D. dissertation, Universität Wien, Wien.

Gunz P, Mitteroecker P, Bookstein FL. 2005. Semilandmarks in three dimensions. In: Slice DE, editor. Modern morphometrics in physical anthropology. New York: Kluwer Academic / Plenum Publishers. p 73-98.

Guy F, Brunet M, Schmittbuhl M, Viriot L. 2003. New approaches in hominoid taxonomy: Morphometrics. Am J Phys Anthropol 121:198-218.

Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4:9pp.

Harvati K. 2003. The Neanderthal taxonomic position: models of intra- and inter-specific craniofacial variation. J Hum Evol 44:107-132.

Harvati K, Frost SR, McNulty KP. 2004. Neanderthal taxonomy reconsidered: implications of 3D primate models of intra- and interspecific differences. Proc Natl Acad Sci U S A 101:1147-1152.

Häusler M, Schmid P. 1995. Comparison of the pelves of Sts 14 and AL 288-1: implications for birth and sexual dimorphism in australopithecines. J Hum Evol 29:363-383.

Hill WCO. 1969. The nomenclature, taxonomy and distribution of chimpanzees. In: Bourne GH, editor. The chimpanzee: a series of volumes on the chimpanzee. Volume 1: Anatomy, behavior, and diseases of chimpanzees. New York: S. Karger. p 22-49.

Holland TD. 1986. Sex determination of fragmentary crania by analysis of the cranial base. Am J Phys Anthropol 70:203-208.

Holloway RL. 1981. The endocast of Omo juvenile L388y-6 hominid: gracile or robust Australopithecus? Am J Phys Anthropol 54:109-118.

Howell FC. 1978. Hominidae. In: Maglio VJ, Cooke HBS, editors. Evolution of African mammals. Cambridge, Mass.: Harvard University Press. p 154-248.

Hughes AR, Tobias PV. 1977. A fossil skull probably of the genus *Homo* from Sterkfontein, Transvaal. Nature 265:310-312.

Immersion Corporation. 1998. Microscribe 3D user's guide & setup instructions. San Jose, CA: Immersion Corporation.

Jackson DA, Somers KM, Harvey HH. 1989. Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence? The American Naturalist 133:436-453.

Johanson DC. 1989. The current status of *Australopithecus*. In: Giacobini G, editor. Hominidae: Proceedings of the 2nd international congress of human paleontology. Milan: Jaca Book. p 77-96.

Jolicoeur P. 1963. The multivariate generalization of the allometry equation. Biometrics 19:497-499.

Jungers WJ, Falsetti AB, Wall CE. 1995. Shape, relative size, and size-adjustments in morphometrics. Yrbk Phys Anthropol 38:137-161.

Keith A. 1925a. The new missing link. British Medical Journal 1:325-326.

Keith A. 1925b. The Taungs skull. Nature 116:11.

Keith A. 1947. Australopithecinae or Dartians. Nature 159:377.

Kelley J. 1993. Taxonomic implications of sexual dimorphism in Lufengpithecus. In: Kimbel WH, Martin LB, editors. Species, species concepts, and primate evolution. New York: Plenum Press. p 429-458.

Kelley J, Plavcan JM. 1998. A simulation test of hominoid species number at Lufeng, China: implications for the use of the coefficient of variation in paleotaxonomy. J Hum Evol 35:577-596.

Kenkel NC, Booth T. 1987. A comparison of presence-absence resemblance coefficients for use in biogeographical studies. Coenoses 2:25-30.

Keyser AW. 2000. The Drimolen skull: the most complete australopithecine cranium and mandible to date. S Afr J Sci 96:189-193.

Kimbel WH. 1984. Variation in the pattern of cranial venous sinuses and hominid phylogeny. Am J Phys Anthropol 63:243-263.

Kimbel WH. 1990. The radiator hypothesis: a theory in “vein”. Behav Brain Sci 13:361-362.

Kimbel WH, Martin LB (eds.) 1993. Species, species concepts, and primate evolution. New York: Plenum Press.

Kimbel WH, Rak Y. 1993. The importance of species taxa in paleoanthropology and an argument for the phylogenetic concept of the species category. In: Kimbel WH, Martin LB, editors. Species, species concepts, and primate evolution. New York: Plenum Press. p 461-484.

Kimbel WH, Rak Y, Johanson DC. 2004. The skull of *Australopithecus afarensis*. Oxford: Oxford University Press.

Kimbel WH, White TD. 1988. Variation, sexual dimorphism, and the taxonomy of *Australopithecus*. In: Grine FE, editor. Evolutionary history of the "robust" australopithecines. New York: Aldine de Gruyter. p 175-192.

Kimbel WH, White TD, Johanson DC. 1984. Cranial morphology of *Australopithecus afarensis*: A comparative study based on a composite reconstruction of the adult skull. Am J Phys Anthropol 64:337-388.

Kimbel WH, White TD, Johanson DC. 1988. Implications of KNM-WT 17000 for the evolution of "robust" *Australopithecus*. In: Grine FE, editor. Evolutionary history of the "robust" australopithecines. New York: Aldine de Gruyter. p 259-268.

Klahr D. 1969. A Monte Carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure. Psychometrika 34:319-330.

Klein RG. 2009. The human career: human biological and cultural origins, 3rd ed. Chicago: University of Chicago Press.

Kramer A, Donnelly SM, Kidder JH, Ousley SD, Olah SM. 1995. Craniometric variation in large-bodied hominoids: testing the single-species hypothesis for *Homo habilis*. J Hum Evol 29:443-462.

Kruskal JB. 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29:1-27.

Kruskal JB. 1964b. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29:115-129.

Kuman K, Clarke RJ. 2000. Stratigraphy, artefact industries and hominid associations for Sterkfontein, Member 5. J Hum Evol 38:827-847.

le Gros Clark WE. 1946. Significance of the Australopithecinae. Nature 157:863-865.

Leakey MG, Spoor F, Brown FH, Gathogo PN, Kiarie C, Leakey L, McDougall I. 2001. New hominin genus from eastern Africa shows diverse middle Pliocene lineages. Nature 410:433-440.

Lee S-H, Wolpoff MH. 2005. Habiline variation: a new approach using STET. Theory Biosci 124:25-40.

Lele S, Richtsmeier JT. 1991. Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data. Am. J. Phys. Anthrop. 86:415-427.

Lele S, Richtsmeier JT. 1995. Euclidean distance matrix analysis: confidence intervals for form and growth differences. Am J Phys Anthropol 98:73-86.

Lieberman DE, Pearson OM, Mobray KM. 2000a. Basicranial influence on overall cranial shape. J Hum Evol 38:291-315.

Lieberman DE, Ross CF, Ravosa MJ. 2000b. The primate cranial base: ontogeny, function, and integration. Yrbk Phys Anthropol 43:117-169.

Lockwood CA. 1997. Variation in the face of *Australopithecus africanus* and other African hominoids. Ph.D. dissertation, University of the Witwatersrand, Johannesburg.

Lockwood CA. 1999. Sexual dimorphism in the face of *Australopithecus africanus*. Am J Phys Anthropol 108:97-127.

Lockwood CA, Kimbel WH, Lynch JM. 2004. Morphometrics and hominoid phylogeny: support for a chimpanzee-human clade and differentiation among great ape subspecies. Proc Natl Acad Sci U S A 101:4356-4360.

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.

Lockwood CA, Menter CG, Moggi-Cecchi J, Keyser AW. 2007. Extended male growth in a fossil hominin species. Science 318:1443-1446.

Lockwood CA, Moggi-Cecchi J. 1998. The systematic position of Stw 183, an adolescent maxilla from Sterkfontein. Am J Phys Anthropol [Suppl] 26:151-152.

Lockwood CA, Richmond BG, Jungers WL, Kimbel WH. 1996. Randomization procedures and sexual dimorphism in *Australopithecus afarensis*. J Hum Evol 31:537-548.

Lockwood CA, Tobias PV. 1999. A large male hominin cranium from Sterkfontein, South Africa, and the status of *Australopithecus africanus*. J Hum Evol 36:637-685.

Lockwood CA, Tobias PV. 2002. Morphology and affinities of new hominin cranial remains from Member 4 of the Sterkfontein formation, Gauteng province, South Africa. J Hum Evol 42:389-450.

Manly BFJ. 1994. Multivariate statistical methods: a primer, 2nd ed. Boca Raton, FL: CRC Press.

Mayr E. 1970. Populations, species, and evolution. Cambridge, MA: Harvard University Press.

McCollum M. 2000. Subnasal morphological variation in fossil hominids: a reassessment based on new observations and recent developmental findings. Am J Phys Anthropol 112:275-283.

McHenry HM, Corruccini RS. 1981. *Pan paniscus* and human evolution. Am J Phys Anthropol 54:355-368.

McKee JK. 1993. Faunal dating of the Taung hominid fossil deposit. J Hum Evol 25:363-376.

McNulty KP. 2004. A geometric morphometric assessment of hominoid crania: conservative African apes and their liberal implications. Annals of Anatomy 186:429-433.

McNulty KP, Frost SR, Strait DS. 2006. Examining affinities of the Taung child by developmental simulation. J Hum Evol 51:274-296.

Miller JA. 1991. Does brain size variability provide evidence of multiple species in *Homo habilis*? Am J Phys Anthropol 84:385-398.

Miller JMA. 2000. Craniofacial variation in *Homo habilis*: an analysis of the evidence for multiple species. Am J Phys Anthropol 112:103-128.

Miller JMA, Albrecht GH, Gelvin BR. 1997. An hierarchical analysis of craniofacial variation in *Homo habilis* using a *Gorilla* analog. Am J Phys Anthropol [Suppl] 24:170.

Miller JMA, Albrecht GH, Gelvin BR. 2004. Craniometric variation in early *Homo* compared to modern gorillas: A population-thinking approach. In: Anapol F, German RZ, Jablonski NG, editors. Shaping primate evolution. Cambridge: Cambridge University Press. p 66-96.

Miller JMA, Gelvin BR. 1993. *Homo habilis* and small robust Australopithecine mandibles: A case of confusion? Am J Phys Anthropol [Suppl] 16:146.

Minchin PR. 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69:89-107.

Moggi-Cecchi J. 1997. Variability in the patterns of dental development of *Australopithecus africanus*: a first approach. J Hum Evol 32:A14.

Moggi-Cecchi J. 2001. Patterns of dental development of *Australopithecus africanus*, with some inferences on their evolution with the origin of the genus *Homo*. In: Tobias PV, Raath MA, Moggi-Cecchi J, Doyle GA, editors. Humanity from African naissance to coming millennia. Firenze and Johannesburg: Firenze University Press and Witwatersrand University Press. p 125-133.

Moggi-Cecchi J, Tobias PV, Beynon AD. 1998. The mixed dentition and associated skull fragments of a juvenile fossil hominid from Sterkfontein, South Africa. Am J Phys Anthropol 106:425-465.

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. 1996. Applied linear statistical models, 4th ed. Boston: WCB McGraw-Hill.

Niewoehner WA. 2005. A geometric morphometric analysis of late Pleistocene human metacarpal 1 base shape. In: Bookstein FL, editor. Modern morphometrics in physical anthropology. New York: Kluwer Academic / Plenum Publishers. p 285-298.

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.

O'Higgins P, Jones N. 2006. Tools for statistical shape analysis. York, UK: Hull York Medical School; url: www.york.ac.uk/res/fme/resources/software.htm.

Olson TR. 1985. Cranial morphology and systematics of Hadar formation hominids and "*Australopithecus*" *africanus*. In: Delson E, editor. Ancestors: The hard evidence. New York: Alan R. Liss. p 102-119.

Partridge TC. 1978. Re-appraisal of lithostratigraphy of Sterkfontein hominid site. Nature 275:282-287.

Partridge TC. 1986. Palaeoecology of the Pliocene and lower Pleistocene hominids of southern Africa: How good is the chronological and palaeoenvironmental evidence? S Afr J Sci 82:81-83.

Petersen TR. 2000. Skeletal robusticity at the ancient Maya sites of K'axob, Belize, and Copán, Honduras. M.A. thesis, University of Houston, Houston, Texas.

Pickering TR, White TD, Toth N. 2000. Brief communication: cutmarks on a plio-pleistocene hominid from Sterkfontein, South Africa. Am J Phys Anthropol 111:579-584.

Picq PG. 1990. L'articulation temporo-mandibulaire des hominidés: biomécanique, allométrie, anatomie comparée et évolution. Paris: Centre National de la Recherche Scientifique.

Pielou EC. 1984. The interpretation of ecological data: a primer on classification and ordination. New York: Wiley.

Plavcan JM. 2003. Scaling relationships between craniofacial sexual dimorphism and body mass dimorphism in primates: implications for the fossil record. Am J Phys Anthropol 120:38-60.

Plavcan JM, Cope DA. 2001. Metric variation and species recognition in the fossil record. Evol Anthropol 10:204-222.

Plavcan JM, Lockwood CA, Kimbel WH, Lague MR, Harmon E. 2005. Sexual dimorphism in *Australopithecus afarensis* revisited: how strong is the case for a human-like pattern of dimorphism? J Hum Evol 48:313-320.

Podani J, Miklós I. 2002. Resemblance coefficients and the horseshoe effect in principal coordinates analysis. Ecology 83:3331-3343.

Pollock DD, Zwickl DJ, McGuire JA, Hillis DM. 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst Biol 51:664-671.

Rayner RJ, Moon BP, Masters JC. 1993. The Makapansgat australopithecine environment. J Hum Evol 24:219-231.

Relethford JH, Blangero J. 1990. Detection of differential gene flow from patterns of quantitative variation. Hum Biol 62:5-25.

Reno PL, Meindl RS, McCollum MA, Lovejoy CO. 2003. Sexual dimorphism in *Australopithecus afarensis* was similar to that of modern humans. Proc Natl Acad Sci U S A 100:9404-9409.

Richmond BG, Jungers WL. 1995. Size variation and sexual dimorphism in *Australopithecus afarensis* and living hominoids. J Hum Evol 29:229-245.

Richtsmeier JT, Deleon VB, Lele SR. 2002. The promise of geometric morphometrics. Yrbk Phys Anthropol 45:63-91.

Robinson JT. 1954. The genera and species of the Australopithecinae. Am J Phys Anthropol 12:181-200.

Rohlf FJ. 2003a. Bias and error in estimates of mean shape in geometric morphometrics. J Hum Evol 44:665-683.

Rohlf FJ. 2003b. tps-Small version 1.20. Ecology and Evolution, SUNY at Stony Brook.

Sarmiento EE. 1993. Sts 19 and the origin of the genus *Homo*. Am J Phys Anthropol [Suppl] 16:173.

SAS Institute. 2007. JMP, version 7.0.2. Cary, North Carolina: SAS Institute, Inc.

Schepers GWH. 1950. The brain casts of the recently discovered *Plesianthropus* skulls. Pretoria: Transvaal Museum Memoir No. 4, pt. 2.

Schultz AH. 1969. The skeleton of the chimpanzee. In: Bourne GH, editor. The chimpanzee: a series of volumes on the chimpanzee. Volume 1: Anatomy, behavior, and diseases of chimpanzees. New York: S. Karger. p 50-103.

Schwarcz HP, Grün R, Tobias PV. 1994. ESR dating studies of the australopithecine site of Sterkfontein, South Africa. J Hum Evol 26:175-181.

Schwartz JH, Tattersall I. 2003. The human fossil record. Volume 2. Craniodental morphology of the genus *Homo* (Africa and Asia). New York: Wiley-Liss.

Schwartz JH, Tattersall I. 2005. The human fossil record. Volume 4. Craniodental morphology of early hominids (genera *Australopithecus*, *Paranthropus*, *Orrorin*) and overview. New York: Wiley-Liss.

Scott RS, Ungar PS, Bergstrom TS, Brown CA, Grine FE, Teaford MF, Walker A. 2005. Dental microwear texture analysis shows within-species diet variability in fossil hominins. 436:693-695.

Seiffert ER, Kappelman J. 2001. Morphometric variation in the hominoid orbital aperture: a case study with implications for the use of variable characters in Miocene catarrhine systematics. J Hum Evol 40:301-318.

Shea BT. 1983a. Allometry and heterochrony in the African apes. Am J Phys Anthropol 62:275-289.

Shea BT. 1983b. Paedomorphosis and neoteny in the pygmy chimpanzee. Science 222:521-522.

Shea BT, Leigh SR, Groves CP. 1993. Multivariate craniometric variation in chimpanzees: implications for species identification. In: Kimbel WH, Martin LB, editors. Species, species concepts, and primate evolution. New York: Plenum Press. p 265-296.

Sherwood RJ, Ward SC, Hill A. 2002. The taxonomic status of the Chemeron temporal (KNM-BC 1). J Hum Evol 42:153-184.

Silverman N, Richmond B, Wood B. 2001. Testing the taxonomic integrity of *Paranthropus boisei sensu stricto*. Am J Phys Anthropol 115:167-178.

Singleton M. 2002. Patterns of cranial shape variation in the papionini (Primates: Cercopithecinae). J Hum Evol 42:547-578.

Singleton M. 2005. Functional shape variation in the cercopithecine masticatory complex. In: Bookstein FL, editor. Modern morphometrics in physical anthropology. New York: Kluwer Academic / Plenum Publishers. p 319-348.

Skelton RR, McHenry HM. 1992. Evolutionary relationships among early hominids. J Hum Evol 23:309-349.

Skelton RR, McHenry HM, Drawhorn GM. 1986. Phylogenetic analysis of early hominids. Curr Anthropol 27:21-43.

Slice DE. 2005a. Modern morphometrics. In: Slice DE, editor. Modern morphometrics in physical anthropology. New York: Kluwer Academic / Plenum Publishers. p 1-45.

Slice DE (ed.) 2005b. Modern morphometrics in physical anthropology. New York: Kluwer Academic / Plenum Publishers.

Smith HF, Grine FE. 2008. Cladistic analysis of early *Homo* crania from Swartkrans and Sterkfontein, South Africa. J Hum Evol 54:684-704.

Smith RJ. 1981. On the definition of variables in studies of primate dental allometry. Am J Phys Anthropol 55:323-329.

Spoor CF. 1993. The comparative morphology and phylogeny of the human bony labyrinth. Ph.D. thesis, Universiteit Utrecht, Utrecht.

StataCorp. 2003. Stata statistical software: Release 8.2. College Station, TX: Stata Corporation.

Strait DS, Grine FE. 2004. Inferring hominoid and early hominoid phylogeny using craniodental characters: The role of fossil taxa. J Hum Evol 47:399-452.

Strait DS, Grine FE, Moniz MA. 1997. A reappraisal of early hominid phylogeny. J Hum Evol 32:17-82.

Strait DS, Weber GW, Neubauer S, Chalk J, Richmond BG, Lucas PW, Spencer MA, Schrein C, Dechow PC, Ross CF, Grosse IR, Wright BW, Constantino P, Wood BA, Lawn B, Hylander WL, Wang Q, Byron C, Slice DE, Smith AL. 2009. The feeding biomechanics and dietary ecology of *Australopithecus africanus*. Proc Natl Acad Sci U S A 106:2124-2129.

Strkalj G, Adams J, Wang Q. 2005. Robert broom's first reconstruction of the adult *Australopithecus*. S Afr J Sci 101:217-218.

Tague RG, Lovejoy CO. 1998. AL 288-1 -- Lucy or Lucifer: gender confusion in the Pliocene. J Hum Evol 35:75-94.

Tattersall I. 1986. Species recognition in human paleontology. J Hum Evol 15:165-175.

Taylor JV, DiBennardo R. 1980. Cranial capacity / cranial base relationships and prediction of vault form: a canonical correlation analysis. Am J Phys Anthropol 53:151-158.

Thackeray JF. 2003. Comparison of hominid fossils attributed to *Australopithecus* and *Homo*. S Afr J Sci 99:241.

Thackeray JF, Dupont LM. 2006. Last recorded occurrence of *Australopithecus africanus* and climatic change in southern Africa. S Afr J Sci 102:50.

Thackeray JF, Mdaka S, Navsa N, Moshau R, Singo S. 2000. Morphometric analyses of conspecific males and females: an exploratory study of extant primate and extinct hominid taxa. S Afr J Sci 96.

Tobias PV. 1967. The cranium and maxillary dentition of *Australopithecus* (*Zinjanthropus*) *boisei*. (Olduvai Gorge, Vol. II). Cambridge: Cambridge University Press.

Tobias PV. 1980. "*Australopithecus afarensis*" and *A. africanus*: critique and an alternative hypothesis. Palaeont Afr 23:1-17.

Tobias PV. 1985. The former Taung cave system in light of contemporary reports and its bearing on the skull's provenance: early deterrents to the acceptance of *Australopithecus*. In: Tobias PV, editor. Hominid evolution: Past, present, and future. New York: Alan R. Liss. p 25-40.

Tobias PV. 1989. Hominid variability, cladistic analysis and the place of *Australopithecus africanus*. In: Giacobini G, editor. Hominidae: Proceedings of the 2nd international congress of human paleontology. Milan: Jaca Book. p 119-128.

Tobias PV. 2001. Conversion in palaeo-anthropology: The role of Robert Broom, Sterkfontein, and other factors in australopithecine acceptance. In: Tobias PV, Raath MA, Moggi-Cecchi J, Doyle GA, editors. Humanity from African naissance to coming millennia. Firenze and Johannesburg: Firenze University Press and Witwatersrand University Press.

Tobias PV, Falk D. 1988. Evidence for a dual pattern of cranial venous sinuses on the endocranial cast of Taung (*Australopithecus africanus*). Am J Phys Anthropol 76:309-312.

Tooby J, DeVore I. 1987. The reconstruction of hominid behavioral evolution through strategic modeling. In: Kinzey WG, editor. The evolution of human behavior: primate models. Albany, NY: SUNY Press. p 183-237.

Trinkaus E. 1990. Cladistics and the hominid fossil record. Am J Phys Anthropol 83:1-11.

Uchida A. 1992. Intra-species variation among the great apes: implications for taxonomy of fossil hominoids. Ph.D. dissertation, Harvard University, Cambridge, MA.

van der Merwe NJ, Thackeray JF, Lee-Thorp JA, Luyt J. 2003. The carbon isotope ecology and diet of *Australopithecus africanus* at Sterkfontein, South Africa. J Hum Evol 44:581-597.

Veroni A, Nikitovic D, Schillachi MA. 2010. Sexual dimorphism of the juvenile basicranium. Am J Phys Anthropol 141:147-151.

Vrba ES. 1982. Biostratigraphy and chronology, based particularly on Bovidae, of southern hominid-associated assemblages: Makapansgat, Sterkfontein, Taung, Kromdraai, Swartkrans; also Elandsfontein (Saldanha), Broken Hill (now Kabwe), and Cave of Hearths. Prétirage, Premier Congrès International de Paléontologie Humaine (CNRS: Nice), pp. 707-752.

Walker A, Leakey RE, Harris JM, Brown FH. 1986. 2.5-Myr *Australopithecus boisei* from west of lake Turkana, Kenya. Nature 322:517-522.

Westergaard B. 1989. A new species concept: cladospecies (and cladosubspecies). Pluralistic realism and biotaxonomy: a radical solution to classification exemplified by hominids. In: Giacobini G, editor. Hominidae: proceedings of the 2nd international congress of human paleontology. Milan: Jaca Book. p 71-74.

White DD, Falk D. 1999. A quantitative reanalysis of the endocast from the juvenile *Paranthropus* specimen L338y-6 from Omo, Ethiopia. Am J Phys Anthropol 110:399-406.

White T, Suwa G, Simpson S, Asfaw B. 2000. Jaws and teeth of Australopithecus afarensis from <aka, Middle Awash, Ethiopia. Am J Phys Anthropol 111:45-68.

White TD. 2003. Early hominids- diversity or distortion? Science 299:1994-1997.

White TD, Folkens PA. 2000. Human osteology, 2nd ed. San Diego: Academic Press.

White TD, Johanson DC, Kimbel WH. 1981. *Australopithecus africanus*: its phylogenetic position reconsidered. S Afr J Sci 77:445-470.

Wiens JJ. 1995. Polymorphic characters in phylogenetic systematics. Syst Biol 44:482-500.

Wolpoff MH. 1971. Competitive exclusion among lower Pleistocene hominids: the single species hypothesis. Man 6:601-614.

Wolpoff MH, Thorne AG, Jelínek J, Zhang Y. 1994. The case for sinking *Homo erectus*. 100 years of *Pithecanthropus* is enough! Cour Forsch Inst Senckenberg 171:341-361.

Wood B. 1991. A palaeontological model for determining the limits of early hominid taxonomic variability. Palaeont Afr 28:71-77.

Wood B. 1994. Hominid paleobiology: Recent achievements and challenges. In: Corruccini RS, Ciochon RL, editors. Integrative paths to the past: paleoanthropological advances in honor of F. Clark Howell. Englewood Cliffs, NJ: Prentice Hall. p 147-165.

Wood B, Lieberman DE. 2001. Craniodental variation in *Paranthropus boisei*: a developmental and functional perspective. Am J Phys Anthropol 116:13-25.

Wood B, Richmond BG. 2000. Human evolution: Taxonomy and paleobiology. J Anat 196:19-60.

Wood BA. 1999. '*Homo rudolfensis*' Alexeev, 1986 -- fact or phantom? J Hum Evol 36:115-118.

Wood BA, Quinney PS. 1996. Assessing the pelvis of AL 288-1. J Hum Evol 31:563-568.

Wood BA, Stack CG. 1980. Does allometry explain the differences between "gracile" and "robust" australopithecines? Am J Phys Anthropol 52:55-62.

Wright SP. 1992. Adjusted *p*-values for simultaneous inference. Biometrics 48:1005-1013.