AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 94:307–325 (1994)

Patterns of Dental Development in *Homo, Australopithecus, Pan,* and *Gorilla*

B. HOLLY SMITH
Museum of Anthropology, University of Michigan, Ann Arbor, Michigan 48109

KEY WORDS Hominidae, Pongidae, Primates, Tooth formation, Discriminant analysis

ABSTRACT Smith ([1986] Nature *323:*327–330) distinguished patterns of development of teeth of juvenile fossil hominids as being "more like humans" or "more like apes" based on statistical similarity to group standards. Here, this *central tendency discrimination* (CTD) is tested for its ability to recognize ape and human patterns of dental development in 789 subadult hominoids. Tooth development of a modern human sample (665 black southern Africans) was scored entirely by an outside investigator; pongid and fossil hominid samples (59 *Pan,* 50 *Gorilla,* and 14 fossil hominids) were scored by the author. The claim of Lampl et al. ([1993] Am. J. Phys. Anthropol. *90:*113–127) that Smith's 1986 method succeeds in only 8% of human cases was not sustained. Figures for overall success of classification (87% humans, 68% apes) mask important effects of teeth sampled and age class. For humans, the power of CTD varied between 53% and 92% depending on the number and kind of teeth available—nearly that of a coin toss when data described only two nearby teeth, but quite successful with more teeth or distant teeth. For apes, only age class affected accuracy: "Infant" apes (M_1 development $\leq$ root cleft complete, unemerged) were usually classed as humans, probably because the present developmental standard for great apes is in substantial error under 3 years of age. "Juvenile" apes ($M_1 \geq$ root 1/4), however, were correctly discriminated in 87% of cases. Overall, CTD can be considered reliable (accuracy of 92% for humans and 88% for apes) when data contrast development of distant dental fields and subjects are juveniles (not infants). Restricting analysis of fossils to specimens satisfying these criteria, patterns of dental development of gracile australopithecines and *Homo habilis* remain classified with African apes. Those of *Homo erectus* and Neanderthals are classified with humans, suggesting that patterns of growth evolved substantially in the Hominidae. To standardize future research, the computer program that operationalizes CTD is now available. © 1994 Wiley-Liss, Inc.

Mann (1968, 1975) proposed that patterns of dental development observed in fossils of juveniles provide a key to the rate of growth of extinct species. The pattern of contrasts between developing teeth in juvenile australopithecines, he argued, resembles humans rather than great apes. He concluded that australopithecines had already achieved the distinct delayed growth characteristic of modern humans. Smith (1986) reconsidered patterns of development in the teeth of fossil hominids and came to the opposite conclusion, that the human rate and pattern of life history had not evolved in early hominids. Mann (1975) emphasized the role of cheek teeth in differentiating human and ape development, whereas Smith (1986) thought a contrast between molar

Received October 21, 1993; accepted February 1, 1994.

© 1994 WILEY-LISS, INC.

and incisor-canine fields held the most information. Differences aside, Mann (1968, 1972, 1975) and Smith (1986, 1989b, 1991a, 1992) agree on the fundamental aspects of the problem: Dynamic aspects of growth can be studied in the fossil record; patterns of development distinguish extant great apes and humans; and growth patterns are important reflections of species growth and aging.

In his earlier studies, Mann was able to describe dental development of four subadult individuals of *Paranthropus robustus* from Swartkrans (Mann, 1968, 1975). By 1986, however, data were available for more than a dozen subadult early hominids from a number of localities, and it seemed appropriate to attempt a numerical analysis of the sample. At the time, comparative data were limited in kind: Information on central tendencies of humans and African great apes was available, but a pool of data on individuals was not. Therefore, I compared individual fossils to central tendencies of humans and of apes, discriminating numerically which species growth standard provided the better fit (Smith, 1986). This technique, here called *central tendency discrimination,* substituted for conventional discriminant analysis. Results suggested that gracile australopithecines and some early *Homo* tended to resemble African apes in pattern of formation of teeth, but that Neanderthals closely resembled modern humans. Perhaps most importantly, the study found evidence that dental development was not a monolith within the Hominidae.

The accuracy of central tendency discrimination was subsequently tested on a small sample of extant taxa (10 humans and 5 great apes), and all 15 cases were correctly classified (Smith, 1989b). Recently, however, Lampl et al. (1993) object that Smith's method misclassifies 92% of a sample of 48 human children as apes or "nonhumans"; if so, it can hardly be trusted to assess fossils.

The purpose of the present study is to test central tendency discrimination (Smith, 1986) on a large sample of extant hominoids and, in addition, to reassess fossil hominids by this method. The present study classifies individuals with humans or African apes, based on similarity of developmental stages of the mandibular teeth I_1-M_2 to group standards, under the following conditions:

1. To strain the method, the human test population (black children from southern Africa) differs from that described by the human growth standards (white children from North America).
2. To maximize independence, an outside investigator, Robert L. Tompkins, scored all 665 human subjects.
3. To balance the test, accuracy is also determined for 109 African apes (scored by the author).
4. To standardize application, the entire operation is computerized and the program is made available to interested researchers.

METHODS AND MATERIALS

The rationale of central tendency discrimination is simple: Developmental scores of an individual's teeth are more likely to appear in synchrony—more likely to agree with each other—when scores are assigned from standards based on the individual's correct species or group. To assess synchrony, stages of tooth formation assigned from human growth standards are compared with those assigned from African ape growth standards for each individual, proceeding in the following manner:

1. Teeth are scored by stage of development (see Fig. 1).
2. A dental age is assigned to each developing tooth of an individual from average standards of *human* development.
3. A second set of dental ages is assigned using average standards of *African ape* development.
4. Ages derived from ape standards are multiplied by a scaling factor to express the score in "human years" to equalize the relative length of growth periods in apes and humans. Maturation of the dentition is achieved at about 20 years in humans and 11 years in apes (Dean and Wood, 1981), so the scaling factor used is 1.8 (20/11 = 1.82).
5. Two standard deviations are computed, one for human ages and one for scaled ape ages; the lower standard de-

viation is taken as the best fit. The case is classified as "more like humans" or "more like apes" in developmental pattern.

Within and between species variance ultimately influences ability of the test to discriminate individuals, but variance is not relevant to steps 1–3, which are simply means of assigning scores that can be analyzed numerically. The method itself stands as a separate framework; the particular growth standards and particular scaling factor that go into central tendency discrimination are replaceable, subject to revision when improvements become available.

The purpose of central tendency discrimination is the same as that of any discriminant analysis: to use knowledge gained from reference groups to sort individuals of unknown affinity.

Growth standards

Central tendency discrimination is quite straightforward in theory, although in practice it depends on growth standards that properly describe central tendency. A real practical handicap is that dental development of great apes is poorly known. Dean and Wood (1981) proposed tentative ages for beginning and ending of crown and root formation in great apes, but could provide no information on substages. Thus, some kind of approximation is required to make any real use of the standard. Smith (1986) used a linear approximation because, using human data from Moorrees et al. (1963), this can be shown to be adequate when growth of three anatomical regions of the tooth (crown, root, and apex) is separately approximated. Moreover, the use of a simple approximation for both genera, it was thought, might "level the playing field" between humans and apes, equalizing the high-quality human growth standards with the tentative ones for apes. In 1986, because of limitations in available data, standards for the mandible were applied to all teeth; this makes teeth from the maxilla appear slightly more deviant from standards than those from the mandible (Smith, 1991a). When more complete standards become available, scores should be assigned from standards for the proper dental arch. A European-derived population (based primarily on the white American sample from Boston and Ohio of Moorrees et al., 1963) was chosen to represent human growth for a purely practical reason: No other major human group has been studied throughout the entire juvenile age range in a manner that is statistically adequate (Smith, 1991b). Mid-sex values were used throughout so that accuracy of classification would be equivalent for males and females.

Computerizing the method

In 1986, numeric scores were assigned to each developing tooth by hand from a table, hand multiplied, and individually calculated—a process difficult to repeat without error for five subjects, much less 500. A computer program is now available to automate the process.

The program SMITH86 accepts data scored in fractional stages of completion of crown, root, or apex in the form of continuous scores in which the region (crown, root, or apex) forming is assigned the value that best describes it from 1% to 100% complete. If original data are in the form of ranked stages of development, they are first converted as in Figure 1. The stage Apex Complete (A 100) is ignored by the program because no age can be assigned to a tooth that has finished its development. Ranked scoring systems with 13 or 14 stages (Moorrees et al., 1963) are strongly preferred to those with few stages (e.g., Demirjian et al., 1973) in the present context; a major drop in precision and discriminating power is predicted for the latter.

Given data on formation stages of teeth, the program assigns dental ages and graphs each subject on human and on ape standards of growth, calculates relevant statistics, and names the growth standard that produces the best fit to the data.

The computer operation matches the original hand tabulations quite well, allowing for rounding effects. It does not, however, allow any use of ages based on eruption or emergence of teeth. Smith (1986) assigned a dental age to reflect recent emergence of a tooth in a few cases, with one exception, all involving the first molar. Footnotes gave the

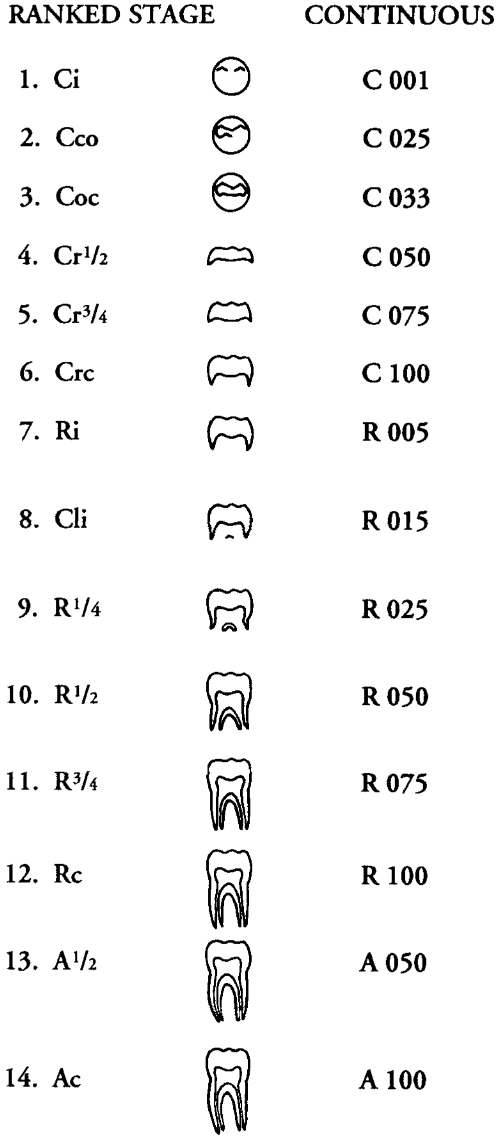

Fig. 1. At left, 14 ranked stages of tooth formation redrawn from Moorrees et al. (1963), a system in which three anatomical regions of the tooth are recognized: C, crown; R, root; and A, apex. At right, equivalents in the continuous style used in the present study. Continuous scores can be assigned directly in any amount from 1% to 100% for the region forming or they can be converted from ranked stages as shown. The program SMITH86 accepts scores in the continuous style. See Moorrees et al. (1963) for further illustrations.

alternative score based on tooth formation if available, and this was noted as problematic. Substitution of emergence age was eliminated in further work (Smith, 1989b). Mixing scores from tooth formation and emergence make the system difficult to apply and test in a standardized manner and, as such, should be eliminated. The program SMITH86 makes no use of data on tooth emergence.

For third molars, the program accepts data but offers a choice of including or excluding them in calculations. Third molars are excluded here because no data for third molars were available to Smith (1986). In practice, data for third molars are uncommon; moreover, substantial geographic differences in human third molar development make these teeth problematic (Fanning, 1962).

Test samples

A difference in age distributions of human and ape samples complicates explanation somewhat: The ape sample includes a substantial portion of very young individuals, whereas the human sample is almost entirely older children. To make statistical comparisons more valid, human and ape samples are each divided into infant and juvenile age classes. The classic anatomical division between infancy and the juvenile period is emergence of the first permanent molar (Schultz, 1960), an event that occurs at about 3 years of age in chimpanzees (Kuykendall et al., 1992) and at about 5½–6 years of age in human populations (Dahlberg and Menegaz-Bock, 1958). Because observations on emergence are not available for the human sample used here, however, age classes must be recognized by tooth formation rather than emergence. In the present study, "infants" are defined as individuals with first permanent molar development ≤stage 8 (root cleft complete, see Fig. 1); "juveniles" are those with first molar development ≥stage 9 (root ¼ complete). This cut was chosen because stage 8 is the last one in which M_1 was always unemerged in the great ape sample. When cut at stages 8/9, 30% of the 109 great apes are infants, whereas only 1% of the 665 humans are (Table 1). Judging from the work of Kuykendall et al. (1992) on chimpanzees and Moorrees et al. (1963) on humans, we might expect

TABLE 1. Test samples by age class

	Infants $M_1 \leq$ stage 8[1]	Juveniles $M_1 \geq$ stage 9	Total
Homo sapiens[2]	6	659	665
Pan troglodytes	16	43	59
Gorilla gorilla	17	33	50
Total	39	735	774

[1] Stages as in Figure 1.
[2] Normal black children from southern Africa.

that the infant category includes, approximately, apes less than about 2½ years of age of humans less than about 4½ years of age.

Humans are presented by a cross-sectional sample of 665 normal black children from southern Africa. Contrast between origin of the test sample (African black) and the growth standards used by the method (North American white) is good for the purposes of a test because such heterogeneity must strain the method. Radiographs were made on living subjects for other purposes and data were collected for another study (Tompkins, 1991). Of the 665 children, 327 are male and 337 are female, and one subject is missing the datum for gender.

African great apes, including 59 *Pan troglodytes* and 50 *Gorilla gorilla,* come from collections in the United Kingdom, from the British Museum (Natural History) and the Powell Cotton Museum. Most subjects were wild-shot. Radiographs of the dried skulls were made by M.C. Dean and formed part of the original sample used to construct the Dean and Wood (1981) standards of great ape dental development. Thus, apes are being tested against standards that should described them in central tendency. In the great ape sample, 35 are probable males, 25 are probable females, and 49 are of unassigned gender.

For the 774 humans and great apes, radiographs of the mandible were observed and stages of formation of the teeth I_1-M_2 were scored by the system of Moorrees et al. (1963) as modified by Anderson et al. (1976) to include 14 stages for each tooth class. Ranked scores were then converted to continuous equivalents for use in the computer program (Fig. 1). Robert L. Tompkins scored all 665 humans; the author scored the 109 great apes.

Teeth of fossil hominids were scored by the author directly in the continuous system. Nine specimens were seen in original radiographs and/or as original specimens, although a few observations were supplemented from Dean (1987); KNM-WT 15000 was seen in sharp casts and original radiographs. Data for Ehringsdorf, Zhoukoudian B-I, Teshik Tash, and Taung were obtained entirely from drawings, published radiographs, or values in the literature (Virchow, 1920; Weidenreich, 1936, 1937; Gremiatski and Nesturkh, 1949; Conroy and Vannier, 1987). Table 2 displays raw data for fossil hominids.

RESULTS

Table 3 displays numerical results from central tendency discrimination for sample cases: two juveniles randomly selected from those with complete data for each genus (*Gorilla* cases are the only two such cases available). By chance, all six draws were male. Figures 2–4 show the graphic output that accompanies three of these subjects. Case 9 of Table 3 (Fig. 2) is a human child whose teeth were assigned dental ages according to two models, human (first line in Table 3; top graph in Fig. 2) and African ape (second line in Table 3; bottom graph in Fig. 2). The human growth model gave dental ages ranging from 6.6 to 7.8 years for the child's seven teeth, with a mean of 7.21 years. An ape model gave a wider range of ages, from 5.2 to 8.3 years of age, with a mean of 6.37 years. The point of the analysis is not the means, however, but the consistency of dental ages of individual teeth within models. Consistency is good under the human model, as shown by the low standard deviation of dental ages (0.42 years). Consistency is poorer when the ages are assigned under the African ape model, with an absolute deviation of 1.03 years, a deviation that for apes is comparable to a standard deviation of 1.03 × 1.8 = 1.86 human years. Scaling African ape ages by 1.8 to express deviations in human years is essentially the same as stretching the African ape standards out visually to equalize lengths of juvenile periods in the two genera, as in Figure 2. All individuals in Table 3 and Figures 2–4 were correctly dis-

TABLE 2. Raw data from fossil hominids studied here[1] in the continuous style accepted by the program SMITH86: Percent of crown (C), root (R), or apex (A) complete for each tooth

	Percent of tooth formed[2]							
Specimen[3]	I1	I2	C	P3	P4	M1	M2	M2
Gibraltar 2	R 025	C 100	C 100	C 067	C 033	R 025	C 050	—
Teshik Tash	A 100	A 100	R 067	R 050	R 033	—	R 041	—
Ehringsdorf	A 100	A 100	A 050	—	R 075	—	R 100	R 005
Zhoukoudian B-I[4]	A 090	A 050	R 033	R 020	R 005	A 050	C 095	—
KNM-WT 15000 max	A 100	A 025	R 067	R 067	R 075	—	R 067	C 067
KNM-WT 15000	A 100	A 100	R 075	R 058	R 058	A 100	R 050	—
KNM-ER 820	R 100[4]	R 075	R 025	R 025	R 020	A 050	R 005	—
KNM-ER 1507	—	—	C 100	R 005	C 095	R 100	C 095	—
KNM-ER 1590 max	R 075	—	C 100	R 033	R 033	A 050	C 100	—
Stw 151 max	R 067	R 050	C 095	C 100	—	R 100	C 100	—
Stw 151	R 075	—	C 100	C 095	C 100	R 100	C 100	—
Taung	C 095	C 095	C 075	C 075	C 75	R 075	C 050	—
LH 2	C 100	—	C 075	C 067	—	R 033	—	—
LH 3 max	R 005	C 100	C 067	C 075	C 095	R 067	—	—
LH 3	—	C 100	C 075	C 075	C 075	R 067	—	—
LH 6 max	—	R 030	C 095	R 005	—	R 075	—	—
SK 63	R 025	R 025	C 095	—	C 095	R 040	C 067	—

[1] See Smith (1989b) for additional specimens.

[2] Assigned directly in continuous system (Fig. 1). Percentages should not be taken to indicate great precision: C 095 = crown almost complete; R 005 = root initiated; R 058 = root 1/2–2/3, etc.

[3] KNM, Kenya National Museums; Stw, Sterkfontein Witwatersrand; LH, Laetoli hominid; SK, Swartkrans; max, maxilla (mandibles unspecified).

[4] Corrected from Smith (1989b) after discovering additional views (Zhoukoudian B–I in Weidenreich, 1936, 1937) or seeing original radiographs (KNM-ER 820).

TABLE 3. Numeric analyses of two cases randomly selected from juveniles with complete data for each extant species

Case no. gender	Model	N	Age μ h.y. a.y.	Age σ h.y. h.y.[1]	Age score (human yrs/ape yrs) I1	I2	C	P3	P4	M1	M2	Best fit
Homo sapiens												
9 ♂	Human	7	7.21	0.42	7.4	7.6	7.1	6.6	7.8	7.3	6.7	Human
	Ape	7	6.37	1.86	7.0	6.7	8.3	5.8	6.0	5.7	5.2	
499 ♂	Human	7	4.95	0.53	4.4	4.8	4.3	5.2	5.5	4.8	5.7	Human
	Ape	7	4.64	1.03	4.3	4.5	5.8	5.0	4.5	4.1	4.4	
Pan troglodytes												
25 ♂	Human	7	6.67	1.99	6.0	6.7	3.1	6.1	7.2	8.1	9.5	Ape
	Ape	7	5.73	1.27	5.8	5.9	4.4	5.5	5.7	6.3	6.6	
38 ♂	Human	7	6.03	1.77	5.2	4.8	3.1	6.1	7.8	7.3	8.0	Ape
	Ape	7	5.26	1.18	5.0	4.5	4.4	5.5	6.0	5.7	5.8	
Gorilla gorilla												
117 ♂	Human	7	6.60	2.02	6.0	5.8	3.1	6.1	9.2	8.1	8.0	Ape
	Ape	7	5.66	1.39	5.8	5.2	4.4	5.5	6.8	6.3	5.8	
127 ♂	Human	7	5.02	1.79	3.8	3.9	2.3	5.2	6.3	7.3	6.4	Ape
	Ape	7	4.50	1.70	3.6	3.7	3.3	5.0	5.2	5.7	5.0	

Dental ages are interpolated from human and African ape standards of dental development, and best fit to standard (ape/human) is shown.

[1] Note that standard deviations for ape dental ages have been multiplied by 1.8 to express them in human years (h.y.); other ages read from normal scale for group (noted as a.y. for ape years and h.y. for human years); see also Figures 2–5.

criminated by the SMITH86 program, as is typical for juveniles with complete data (see below).

With 774 subjects is becomes possible to look for effects of number of teeth, combination of teeth, age, and gender on percent classified correctly. Tables below always show the effect of number of teeth, but, for the sake of clarity, results are broken down for other factors only if they can be shown to be important.

For human subjects overall (Table 4), 86.6% of the 665 were classified with human as opposed to ape standards, but success varied markedly with sampling conditions. If more than two teeth were available, success rose to 89.8%. The kind of teeth available, however, was more important than the

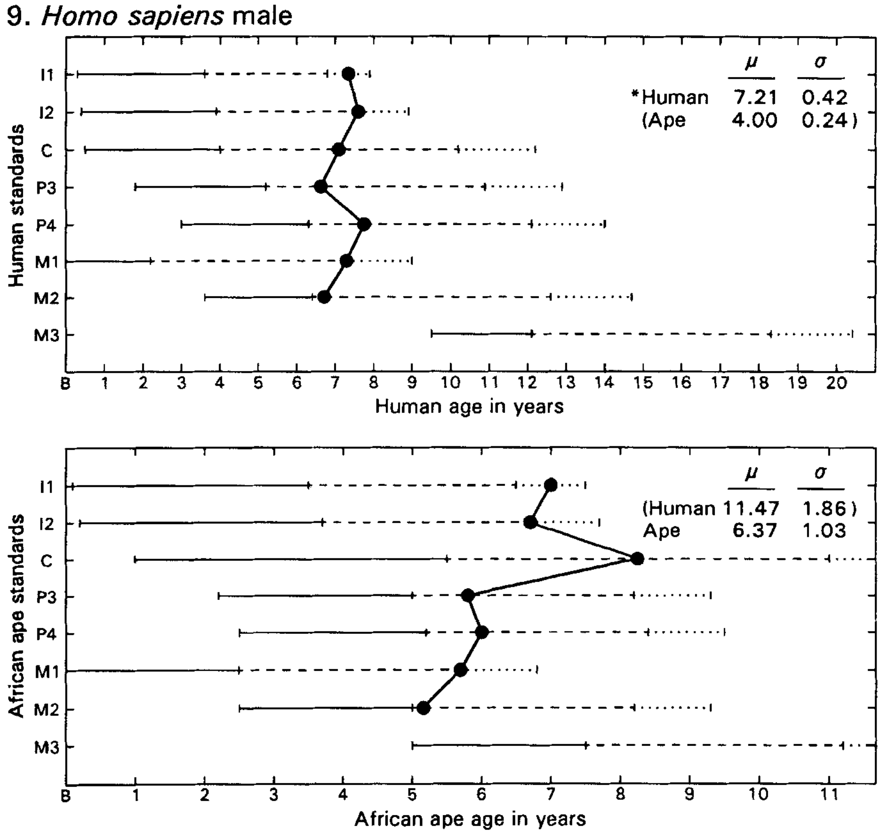

Fig. 2. Mandibular teeth of human child (case no. 9) plotted on human (top) and African ape (bottom) standards of permanent tooth formation; case selected randomly from subjects with complete data (see Table 1). Charts are modified from Dean and Wood (1981). Solid lines, crown formation; dashed lines, root formation; dotted lines, apex formation. The ape standard is magnified by a factor of 1.8 to equalize periods of tooth formation. The child's teeth show greatest synchrony when plotted on the human chart (that is, plotted points come nearer a vertical line); they are comparatively out of synchrony on the ape chart. For computations, values can be scaled in either "human years" (the birth-20 scale) or "ape years" (the birth-11 scale); mean and standard deviation are given in both scales above. ( ) indicates scale is taken from the opposing graph; * denotes standard with best numeric fit to data.

number: For humans, success reached 91.6% when data contrasted molar and incisor-canine fields, and the method appeared to work about as well with two teeth as with three, four, five, six, or even seven teeth; even N = 2 produced a success of 91%. Without a "contrast," success fell to approximately ¾ at N = 3–4 teeth, and approached a coin toss at N = 2 teeth. The small number of human infants precludes an effective test of age class (infant versus juvenile); however, five of the six infants (83.3%) were correctly classified. Gender had no effect whatsoever (success of 86.2% for males and 86.9% for females).

One subtle effect of within-sample age underlies Table 4. As children mature, teeth complete formation and drop out of the study, so that age and number of teeth available are correlated. In the present human sample, the rank-order correlation between number of teeth with data present and stage of M2 development is substantial, rho = −0.80. This "drop-out" effect produces the category with the worst outcome in Table 4—that of two teeth, no "contrast," and 53%

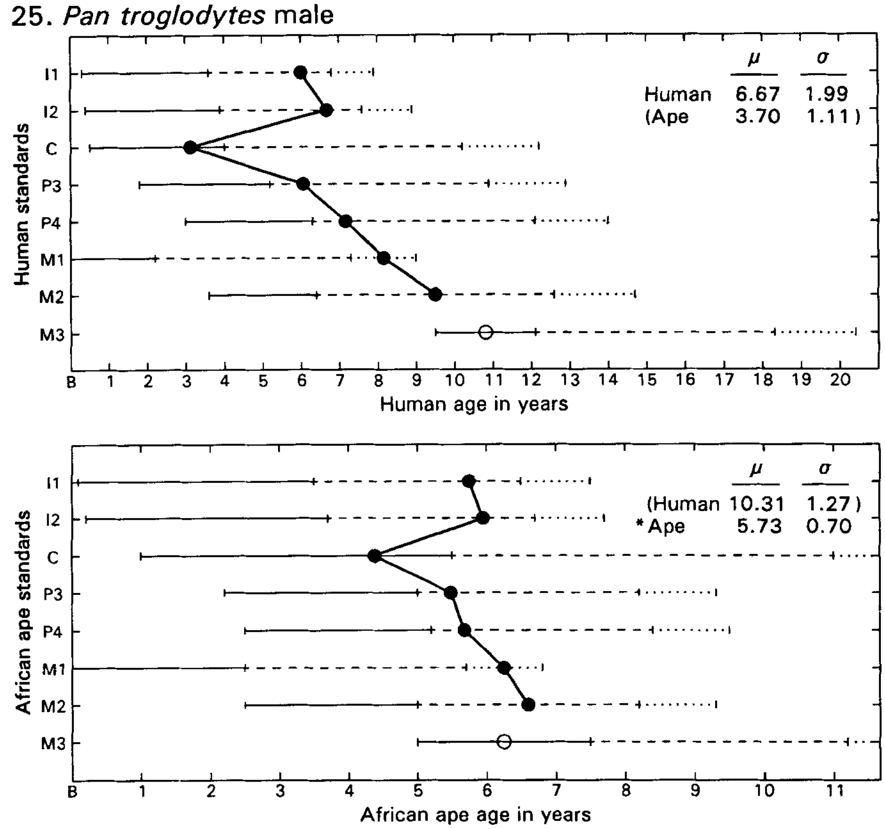

Fig. 3. Mandibular teeth of *Pan troglodytes* (case no. 25) plotted on human (top) and African ape (bottom) standards of permanent tooth formation; case selected randomly from juveniles with complete data. Definitions as in Figure 2. Third molar datum is displayed, but not included in calculations. The juvenile's teeth show greater synchrony when plotted on the ape chart, where points show relatively less dispersion. This individual is in marked discordance with the human standard, where dental ages range from 3.1 to 9.5 years; fit on the ape chart is better, although imperfect. Scale as in Figure 2; ( ) indicates numeric scale is taken from the opposing graph; * denotes standard with best numeric fit to data.

success. In this category, the two teeth present are not random pairs: 53 of the 58 are cases with data only for P_4 and M_2, the last two teeth to complete formation in humans before third molars.

Central tendency discrimination was less successful at grouping apes with their proper standard than humans with theirs: Only 67.9% of African apes were correctly classed overall (Table 5). Age group, however, shows a striking effect on percent correctly classed: Most errors occurred in the infant class where the program called 3/4 infants "human." Results for infants (24.2% correct) are in stark contrast to those for juveniles (86.8% correct), a highly significant difference ($P < 0.001$, chi-squared test). No other factor could be shown to affect accuracy when age was controlled. By gender, success was 78% for 32 probable male and 95% for 20 probable female juveniles, but the difference was not significant ($P = 0.213$). Similarly, by genus, a 95% success for 43 *Pan* vs. 79% for 33 *Gorilla* juveniles was not a significant difference ($P = 0.140$).

Apes were less susceptible than humans to the drop-out effect at older ages because in great apes the canine keeps forming throughout the subadult period. A datum for the canine was almost always present in the

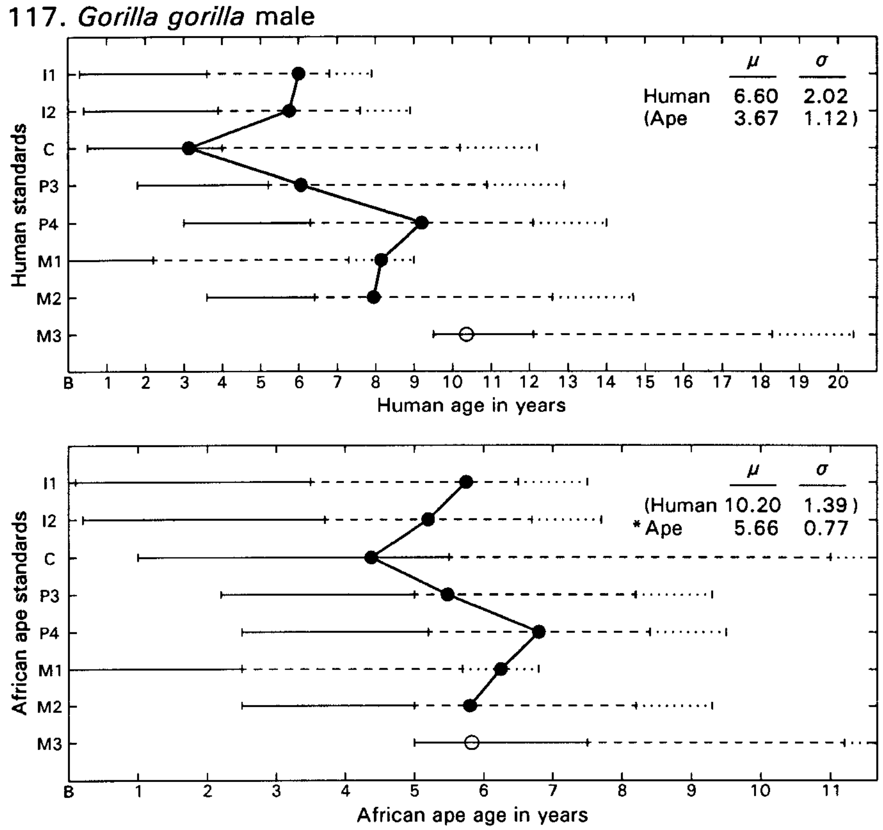

Fig. 4. Mandibular teeth of *Gorilla gorilla* (case no. 117) plotted on human (**top**) and African ape (**bottom**) standards of permanent tooth formation; the first of only two cases with complete data for juvenile gorillas (Table 2). Definitions as in Figure 2. Third molar datum is displayed, but not included in calculations. Teeth are in marked discordance with the human standard where dental ages range from 3.1 to 9.2 years; fit on the ape chart is better, although imperfect. Scale as in Figure 2; ( ) indicates numeric scale is taken from the opposing graph; * denotes standard with best numeric fit to data.

ape data set. Indeed, almost all subjects possessed data for both molar and incisor-canine fields, so distant dental fields could be contrasted. Simulations of missing data are not included because they would not recreate real patterns of missing data in great apes.

All in all, central tendency discrimination works best (an overall accuracy approximating 90% for both humans and apes) when two sampling criteria are met: Incisor-canine and molar fields are contrasted and subjects are juveniles, not infants. Under these sampling conditions, juvenile humans are classed as most similar to human standards in 91.7% of cases and juvenile African apes are correctly classified in 87.8% of cases (Table 6).

Criteria for sampling adequacy developed on extant taxa are used to constrain cases of fossil hominids analyzed in Table 7. To satisfy the new criteria, eight specimens included in 1986 (two gracile and six robust australopithecines) must be discarded. However, more data and more individuals have been added since 1986, and, in all, 14 specimens satisfy new criteria. As shown in Table 7, the analysis classified "archaic" *Homo sapiens* as human, along with most specimens of *Homo erectus* (Fig. 5). In contrast, the single *Homo habilis* specimen is classified with African apes, as are most

TABLE 4. Percent of human children correctly classified, sorted by factors (I–IV) that affect classification

	I. Number of teeth						
	Two	Three	Four	Five	Six	Seven	Total
II. Field contrast[1]							
Contrast							
Correct %	90.9	93.9	97.2	86.5	82.1	93.6	**91.6**
Incorrect %	9.1	6.1	2.8	13.5	17.9	6.4	8.4
N	11	33	178	111	78	125	536
No contrast							
Correct 1%	53.4	76.8	73.3				**65.9**
Incorrect %	46.6	23.2	26.7				34.1
N	58	56	15				129
III. Age: Cannot test[2]							
IV. Gender: no effect							
Total sample							
Correct %	59.4	83.1	95.3	86.5	82.1	93.6	**86.6**
Incorrect %	40.6	16.9	4.7	13.5	17.9	6.4	13.4
N	69	89	193	111	78	125	665

[1] Contrast of molar and incisor-canine field: Data present for at least one of I1, I2, or C, and one of M1 or M2. At five or more teeth, such a contrast is always present.

[2] N too small, but of the six infants, five (83.3%) were correctly classed.

TABLE 5. Percent of great apes correctly classified sorted by factors (I–IV) that affect classification

	I. Number of teeth						
	Two	Three	Four	Five	Six	Seven	Total
II. Field contrast: not tested[1]							
III. Age							
Juveniles							
Correct %	80.0	100.0	100.0	92.9	88.9	77.3	**86.8**
Incorrect %	20.0	0.0	0.0	7.1	11.1	22.7	13.2
N	5	2	6	14	27	22	76
Infants							
Correct %	100.0	83.3	0.	16.7	33.3	0.0	**24.2**
Incorrect %	0	16.7	100.0	83.3	66.7	100.0	75.8
N	1	6	9	6	3	8	33
IV. Gender: No effect							
Total sample							
Correct %	83.3	87.5	40.0	70.0	83.3	56.7	**67.9**
Incorrect %	16.7	12.5	60.0	30.0	16.7	43.3	32.1
N	6	8	15	20	30	30	109

[1] Contrast of molar and incisor-canine field: Data present for at least one of I1, I2, or C, and one M1 or M2. At five or more teeth, such a contrast is always present. For apes, 107 of 109 had contrast.

gracile *Australopithecus*. Only one robust australopithecine remains after discarding young or incomplete specimens. The remaining case, SK 63, however, is a solid one that continues to be grouped with humans.

DISCUSSION

Comparison with other studies

The Smith (1986) method of discriminating human and ape patterns of tooth formation (here called *central tendency discrimination*) was tested on 10 humans and 5 great apes with 100% success in Smith (1989b). Although the latter study used a slightly different calibration of rate of root growth for African apes, use of the original calibration also gives 100% success. Human subjects in that study were diverse, including recent Canadians and archaeological specimens from Hudson Bay and the Middle East; great apes were both captive and wild. All 15 subjects had data for five or more teeth, providing a contrast between distant dental fields.

Recently, Lampl et al. (1993) attribute a substantial error to Smith's 1986 method of classifying patterns of dental development: The abstract claims an error of 92%, al-

TABLE 6. Summary of percent of cases correctly classified under adequate sampling conditions (dental fields "contrast," juvenile age) in present study

	Number of teeth						
	Two	Three	Four	Five	Six	Seven	Total
Humans							
Correct %	90.0	93.9	97.2	86.4	81.6	94.3	91.7
Incorrect %	9.1	6.1	2.8	13.6	18.4	5.7	8.3
N	11	33	178	110	76	122	530
African apes							
Correct %	100.0	100.0	100.0	92.9	88.9	77.3	87.8
Incorrect %	0.0	0.0	0.0	7.1	11.1	22.7	12.2
N	3	2	6	14	27	22	74

though a count of their Tables 1–2 gives 56% error/44% success.[1] In the present study of humans, when sampling contrasted dental fields as originally recommended in Smith (1986), *success*—not failure—was 92%. The Smith (1986) method is indeed poor at sorting patterns of infants, but this is not the source of the Lamp et al. result because that study includes only four infant cases.

It is not possible to reconcile results of the present study with those of Lampl et al. (1993) because their specimens are not available for study, but a few comments can be made. There are at least four sources of variation between studies: chance, population, sampling conditions, and methods. In the present case, the first three appear unimportant. (1) Chance can be ruled out by simple probability because 92% failure can hardly be drawn from the same universe as 92% success for the sample sizes at issue ($P < 10^{-8}$). (2) Population: The Middle Eastern and Native American subjects seen by Lampl et al. (1993) should be far easier—not harder—to classify than the African black subjects considered here (Smith 1991b; Smith, 1994). Further, Smith (1989b) succeeded with Middle Eastern and Native American subjects. (3) Sampling was good for the majority of the Lampl et al. (1993) specimens; most had an adequate representation of teeth per individual and only four infants were included. Limiting Lampl et al. (1993) cases to those strictly comparable to the present study does not materially increase resemblance of their results to the present study (Table 8 below). In their study, classification was poor for 23 mandibles well sampled (52% success), poor for 19 maxillae well sampled (44% success), and poor for 7 cases poorly sampled (no "contrast"/infants; 57% success). Thus, the Lampl et al. (1993) results (Table 8) conflict with those of the present study (Table 6).

The remaining possibility is that disparate results stem from disparate methods, and oddities in scores in Tables 1–2 of Lampl et al. (1993) support this interpretation. For example, in the present study a first incisor at stage 12 ("root complete") is scored 6.5 years under an ape model and 6.8 years under the human one. Thus, for I1 every score of 6.5 (ape) should be paired with a score of 6.8 (human). In Tables 1–2 of Lampl et al. (1993), however, the 6.5 (ape) score is variously paired with human scores of 6.8, 7.3, 7.5, *or* 7.8 years. Multiple scoring appears again for M1 (the ape score of 5.7 is paired with human scores of 7.3, 8.2, and 9.0), and for M2 (ape score of 5.4 appears with both 6.8 and 7.3). Scores for I2 are quite exceptional: The ape score of 6.6 is paired with human scores of 7.8, 8.1, 8.5, and 9.0, and the correct value (7.5) never appears. Such multiple scores appear only when assigned

---

[1] Lampl et al. (1993) never state the percent error of their attempt at central tendency discrimination, but a count of results in their Tables 1–2 puts it at 56%. The 92% error rate featured in the Lampl et al. abstract actually comes from an entirely different procedure, which they incorrectly attribute to Smith (1986). Figure 3 in Smith (1986) displayed data for robust and nonrobust clades of early hominids. In it a dashed reference line served to point out the expected central tendency for humans, but the figure played no part in classifying individuals by pattern. Lampl et al. produce a similar figure but use it to classify individuals as follows: Those adhering perfectly or near-perfectly to the dashed line are "human," and all those deviating in either a plus or minus direction are labelled "ape" or "nonhuman." Given the absurdity of the protocol, it is surprising success was as high as 8%.

TABLE 7. Classification of dental development of fossil hominids for those that satisfy minimal conditions for adequate sampling (dental fields "contrast," juvenile age)

			Age μ	Age σ	Age score (human yrs/ape yrs)							
Specimen[1]	Model	N	h.y. a.y.	h.y. h.y.[2]	I1	I2	C	P3	P4	M1	M2	Best fit
"Archaic" *Homo sapiens*/Neanderthals												
Gibralter 2	Human	7	4.13	0.47	4.4	3.9	4.0	4.1	4.1	3.5	5.0	Human
	Ape	7	4.00	1.34	4.3	3.7	5.5	4.1	3.4	3.3	3.8	
Teshik Tash	Human	4	8.34	0.41	—	—	8.2	8.0	8.2	—	8.9	Human
	Ape	4	7.09	2.53	—	—	9.2	6.6	6.3	—	6.3	
Ehringsdorf	Human	3	11.48	1.01	—	—	11.2	—	10.7	—	12.6	Human
	Ape	3	9.23	4.19	—	—	11.9	—	7.6	—	8.2	
Homo erectus												
Zhou B–I	Human	7	7.06	0.96	7.8	8.3	6.0	6.3	6.6	8.1	6.3	Human
	Ape	7	6.29	1.86	7.4	7.2	7.3	5.6	5.4	6.3	4.9	
KMN-WT 15000 max	Human	5	9.26	1.29	—	7.9	8.2	9.0	10.7	—	10.6	Human
	Ape	5	7.60	1.65	—	6.9	9.2	7.1	7.6	—	7.1	
KNM-WT 15000[3]	Human	4	9.08	0.59	—	—	8.6	8.5	9.7	—	9.5	Human
	Ape	4	7.53	2.53	—	—	9.6	6.9	7.1	—	6.6	
KNM-ER 820	Human	7	6.85	0.80	6.8	6.7	5.6	6.6	7.5	8.1	6.7	Human
	Ape	7	6.05	0.99	6.5	5.9	6.9	5.8	5.8	6.3	5.2	
KNM-ER 1507[4]	Human	5	5.84	1.22	—	—	4.0	5.5	6.1	7.3	6.3	Ape
	Ape	5	5.26	0.60	—	—	5.5	5.2	5.1	5.7	4.9	
Homo habilis												
KNM-ER 1590	Human	6	6.64	1.57	6.0	—	4.0	7.1	8.2	8.1	6.4	Ape
	Ape	6	5.80	0.88	5.8	—	5.5	6.1	6.3	6.3	5.0	
Australopithecus africanus												
Taung	Human	7	4.45	1.09	3.4	3.7	3.1	4.3	5.5	6.0	5.0	
	Ape	7	4.10	1.03	3.3	3.5	4.4	4.3	4.5	4.9	3.8	
STW 151 max	Human	6	5.70	1.17	5.7	5.8	3.8	5.2	—	7.3	6.4	Ape
	Ape	6	5.28	0.50	5.5	5.2	5.3	5.0	—	5.7	5.0	
STW 151	Human	6	5.84	1.16	6.0	—	4.0	5.0	6.3	7.3	6.4	Ape
	Ape	6	5.34	0.67	5.8	—	5.5	4.9	5.2	5.7	5.0	
Australopithecus afarensis												
LH 2	Human	4	3.67	0.41	3.6	—	3.1	4.1	—	3.9	—	Human
	Ape	4	3.88	0.76	3.5	—	4.4	4.1	—	3.6	—	
LH 3 max	Human	6	4.43	1.23	3.8	3.9	2.8	4.3	6.1	5.6	—	Ape
	Ape	6	4.23	1.00	3.6	3.7	4.0	4.3	5.1	4.6	—	
LH 3	Human	5	4.49	1.06	—	3.9	3.1	4.3	5.5	5.6	—	Ape
	Ape	5	4.31	0.66	—	3.7	4.4	4.3	4.5	4.6	—	
LH 6 max	Human	4	5.09	0.94	—	5.0	3.8	5.5	—	6.0	—	Ape
	Ape	4	4.98	0.54	—	5.3	5.2	—	4.9	—		
Paranthropus robustus												
SK 63	Human	6	4.82	0.86	4.4	4.8	3.8	—	6.1	4.2	5.5	Human
	Ape	6	4.50	1.02	4.3	4.5	5.3	—	5.1	3.8	4.2	

Dental ages are interpolated from human and African ape standards of dental development and best fit to standard (ape/human) is shown.

[1] Zhou, Zhoukoudian; KNM, Kenya National Museums; Stw, Sterkfontein Witwatersrand; LH, Laetoli hominid; max, maxilla (mandibles unspecified).

[2] Note that standard deviations for ape dental ages have been multiplied by 1.8 to express them in human years (h.y.); other ages read from normal scale for group (a.y., ape years).

[3] The linear extrapolation used here to assign scores underestimates age, but this is not critical in pattern assessment. Including M3 and using proper age prediction tables for males gives a mean of 10.5 years for KNM-WT 15000 under a human model (Smith, 1993).

[4] KNM-ER 1507 is *Homo* aff. *erectus* in Wood (1991).

under the human model, never the ape model, increasing discrepancies for human scores. In a similar vein, two scores of 9.0 at I2 are greater than the maximum age assignable for a human score, although no such irregularity occurs under the ape model. Substitution of emergence ages at I2, C, M1, and M2 added yet more variance into the system. It is not surprising that the method fared poorly under such a regime of score assignment.

How different are apes and humans?

For the human population tested here, the power of central tendency discrimination (Smith, 1986) varied between 53% and 92% depending on the number and kind of teeth available—hardly better than a coin toss for

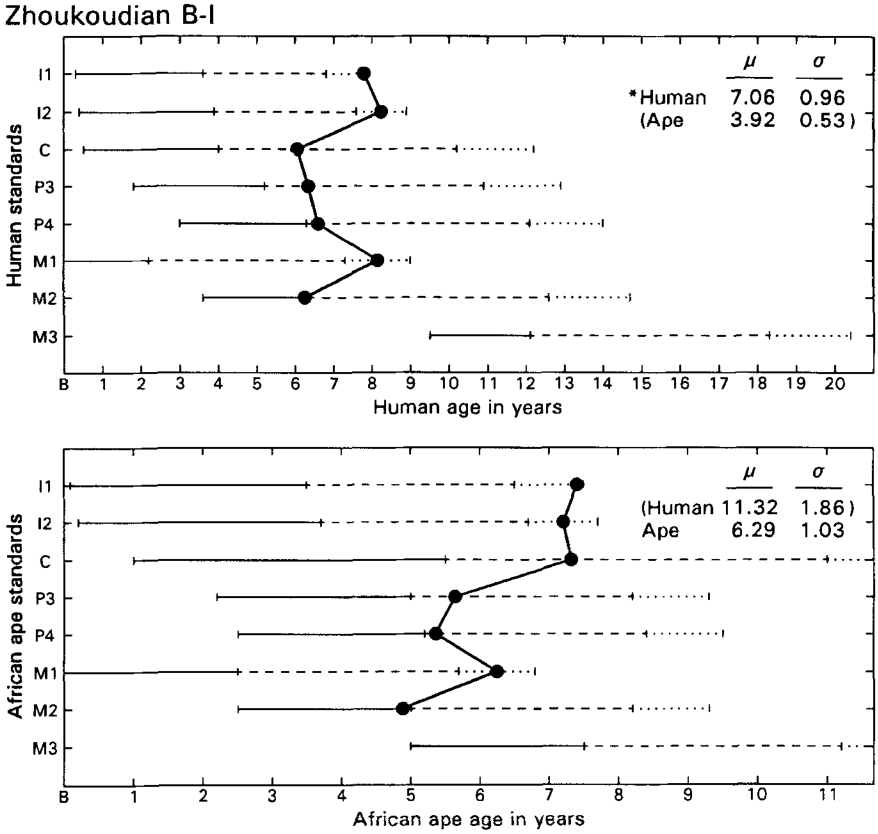

Fig. 5. Mandibular teeth of Zhoukoudian mandible B-I of *Homo erectus* plotted on human (top) and African ape (bottom) standards of permanent tooth formation. Definitions as in Figure 2. Dispersion is relatively less when plotted on human standards of development, although B–I shows a poorer fit than the human child in Figure 2. Plotted on the ape chart, anterior teeth are too advanced for posterior teeth. Scale as in Figure 2. ( ) indicates numeric scale is taken from the opposing graph; * denotes standard with best numeric fit to data.

two nearby teeth, but quite successful with more teeth or distant teeth. A look at a few simple tabulations of developmental scores helps explain why this is the case.

Table 9 indicates the degree of true overlap between apes and humans, comparing a gorilla to a group of humans matched for M_2 development. The choice of tooth and stage to match was based only on the likelihood that data would be available from both human and ape data sets. Six apes and 41 humans had M_2 at developmental stage 6, but to make the comparison manageable, the single most typical ape (a gorilla) was chosen to compare with all 41 humans. Table 9 shows how many of the 41 humans matched the reference gorilla in stages of development of other teeth. As shown, matches were confined to the cheek teeth P_3-M_1, where matches were numerous for small sets of teeth: 31 humans matched the gorilla at two tooth positions and 15 matched at three. One human matched the gorilla over four cheek teeth, but here the resemblance ends. No human child could be found to match the gorilla in developmental stage of more than four teeth, even with 41 potential matches. Moreover, no humans matched at I_1, I_2, or C, suggesting that anterior and posterior teeth are on widely differing developmental tracks in the two genera. All in all, if the gorilla record was mixed in with those of

TABLE 8. Summary of percent of cases correctly classified under adequate sampling conditions (dental fields "contrast," juvenile age)[1,2] *as reported by Lampl et al. (1993) for their human data*

	Number of teeth						
	Two	Three	Four	Five	Six	Seven	Total
Mandible							
Correct %	50.0	50.0	66.7	20.0	60.0	100.0	52.2
Incorrect %	50.0	50.0	33.3	80.0	40.0	0.0	47.8
N	2	4	6	5	5	1	23
Maxilla							
Correct %	—	33.3	25.0	50.0	50.0	66.7	44.4
Incorrect %	—	66.7	75.0	50.0	50.0	33.3	55.6
N	0	3	4	4	4	3	18

Compare with Table 6.

[1] M3 data discarded and case recalculated for comparability in five cases. Only mandibles are strictly comparable with present study, but maxillae are included for comparison.

[2] Cases excluded for sampling considerations include four infants and three cases without "contrast"; success was 57% in this group

TABLE 9. Extent of true overlap between ape and human dental development

Combination of teeth matching (all match at M_2)	Number of human cases matching a reference gorilla (%)[1]	Usefulness of combination for discriminating ape/human
M_2,I_1[2]	0/41 (0)	↑
M_2, I_2	0/41 (0)	High
M_2, C	0/41 (0)	
M_2, M_1, P_4, P_2	1/41 (2)	
M_2, M_1, P_3	2/41 (5)	
M_2, M_1, P_4	4/41 (10)	
M_2, P_4, P_3	11/41 (27)	
M_2, P_3	13/41 (32)	
M_2, M_1	16/41 (39)	
M_2, P_4	20/41 (49)	Low
M_2	41/41(100)	↓

A reference gorilla shares M_2 development of stage 6 with 41 human cases: How many humans show additional matches in development of other teeth?

[1] In stage of dental development for that tooth combination; developmental scores for reference case (gorilla 127) are for II-M2: 7, 6, 4, 6, 6, 12, and 6 (see Fig. 1 and Table 3).

[2] No matches: for I1–c the gorilla scores of 7, 6, and 4 are far from the human range (respectively, 11–14, 9–12, and 6–10.

the 41 humans, the worst teeth to use to pick it out would be M_2 and P_4, where 49% of humans match the gorilla exactly. Any combination involving M_2 and an anterior tooth, however, would instantly pinpoint the gorilla.

The above exercise suggests that identity of scores between humans and apes may be rare when data are fairly complete, at least for individuals in the midst of the juvenile period. Identity is not rare, however, when only a few nearby or adjacent teeth are sampled, at least at the present level of measurement. This may also be the case for infants. At birth, infant apes and humans may share identical scores for permanent teeth, a condition that is approached again as subjects end the juvenile period and all scores converge to 14. No sorting algorithm can succeed 100% of the time when some individuals in different groups have identical scores.

Advantages and disadvantages

Given the patterns in Table 9, discrimination probably ought to be more successful than it is when data are present for five or six teeth of juveniles (as in Table 4). Some portion of error must stem from differences between the particular human population represented in growth standards vs. that used as a test sample; as is, the program would be expected to perform best on European-derived test subjects. Refinements in score assignment (presently done with a very simple linear approximation) should also improve accuracy. Inaccuracies in the growth standard for African apes, however, may be the most important source of error. The Dean and Wood (1981) standard for dental development of great apes was an enormous step forward and is undoubtedly far more right than wrong. Working with individual cases, however, it becomes clear that real apes never fit the ape standard as well as real humans fit the human one (for example, contrast Fig. 2 with Figs. 3 and 4). Graphs of over 100 apes, taken together, suggest that the great ape standard has two main errors of pattern: First, it errs slightly in the direction of being more like humans than real apes actually are, and second, ages of initial appearance of canine and premolars are substantially too late. Fortunately, several investigators are currently contributing new data aimed at improving standards of tooth formation in great apes (Anemone et al., 1991; Beynon et al., 1991; Kuykendall, 1992).

Infants pose some special problems for studies of development, and the failure to classify infants correctly could stem from problems unique to this age group. Possible explanations are (1) the great ape standard is substantially inaccurate under the age three (i.e., it describes no taxon, whereas the human standard at least describes a real hominoid); (2) apes and humans converge in pattern towards infancy so that differences become less than measurement error or become swamped by approximations; or (3) the scaling factor is too large for infants. Of these possible sources of error, only a problem involving use of a single scaling factor would be fundamental to the method.

Despite limitations imposed by quality of current growth standards, central tendency discrimination has some advantages: It can be used when central tendency is known but variance is incompletely understood, a common state in anthropology and paleontology; it is unaffected by scales of measurement that are not uniform with time; it controls for age changes in pattern (always in theory, but in practice, not yet for infants); and a missing datum presents no difficulty (discriminate functions, in comparison, must be recalculated for every combination of available teeth). Moreover, in assessing fossils, the method graphs each individual relative to two reference populations, a great help to interpretation. A more subtle point in favor of the method is that it may be safer to compare fossils to excellent growth standards than to ordinary pools of sampled individuals because distributions of developmental scores are very sensitive to sample age structure. Comparisons to pools of individuals should draw those individuals from equal age distributions to avoid bias (Smith, 1991b, 1992); use of properly calculated growth standards avoids this problem.

Fossil hominids

Smith (1986) sorted a small sample of fossil hominids using the method tested here. One interesting result was that fossils in different taxonomic groups gave different answers. A Neanderthal child displayed close similarity with modern humans, whereas many older fossils resembled great apes. One truly unexpected result was that robust australopithecines appeared closer to humans than gracile australopithecines, an oddity given standard interpretations of phylogeny (e.g., Tobias, 1967). As Smith (1989b, 1991b) pointed out, available juvenile "robusts" were very young, most under 2.5 years of age by any standard. This alone could explain why the method classed them with human standards, given what it does for infant apes. Other factors, however, may also be at work. A number of observers have noted parallels between dental development of robust australopithecines and humans (Broom and Robinson, 1951; Mann, 1975; Smith, 1986; Conroy and Vannier, 1991). Some of these have been dismissed (Bromage, 1987; Grine, 1987; Conroy, 1988; Smith, 1991a) but others have not, and robust australopithecines may show some true parallels with humans. More study of robust australopithecines is in order. An adequate comparative sample of infant humans and correction of great ape standards for infants might also help resolve the conundrum.

The present study finds that the central tendency discrimination of Smith (1986) distinguishes African apes and humans very well when subjects are juveniles and distant rather than adjacent teeth are sampled. A review of specimens of fossil hominids that satisfy sampling criteria, restricting data to tooth formation (not eruption), demonstrates that the results of Smith (1986) still stand: Members of "archaic" *Homo sapiens,* including two Neanderthals, remain classed with humans; *Homo habilis* and gracile australopithecines, on the other hand, tend to be classified with apes. Since 1986 more complete data for *Homo erectus* have greatly increased the resemblance of that species to modern humans. The present study also suggests that most robust australopithecines are too young to be confidently assessed with present tools of analysis.

Why classify?

The point of the analysis of Smith (1986) was never classification alone. As in many applications of discriminant analysis, classi-

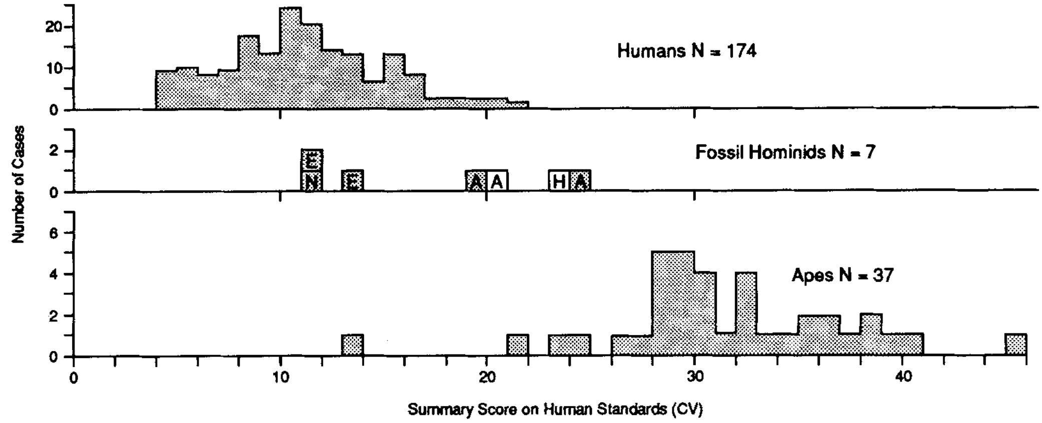

Fig. 6. Distance from North American white patterns of tooth formation expressed by a summary score for each individual (coefficient of variation of human dental age I_1-M_2). High scores reflect highly deviant patterns of tooth formation relative to the reference population. Scores for 174 black children from southern African (**top**) and 37 African apes (**bottom**) are compared to 7 fossil hominids (**middle**). The human distribution is scaled by $1/5$ to equal mass of ape distribution; sample is reduced because nearly complete data were required for inclusion. N, Gibraltar 2 Neanderthal; E, *Homo erectus* Zhoukoudian B-I and KNM-ER 820; A, *Australopithecus africanus* STW 151 and Taung; H, *Homo habilis*. The two maxillae of hominids are left unshaded; all others are mandibles. As expected, the distribution of human scores is closest to the North American standard, whereas the African ape distribution reflects highly deviant patterns. Fossils appear to fall into two groups: Neanderthal and *Homo erectus* with *Homo sapiens*, with *Australopithecus* more distant. The distant value for *Homo habilis* may be partly because it is a maxilla; no data yet available for a mandible.

fication was primarily a heuristic exercise. Rather, the point was to quantify dental development and to learn something about the degree of difference between living humans and apes, and the similarity of extinct species to them. Sorting fossils into bins "like humans" or "like apes" was only a first step towards understanding the adaptations of extinct species. The more crucial issue is where extinct species lie on the continuum between humans and apes; and, indeed, subsequent studies (Smith, 1989b, 1991b, 1993, 1994) emphasize continuous measures of similarity in dental development.

There are many ways to express a pattern of dental development as a continuous variable, and Figure 6 shows one of these (see Simpson et al., 1990, for another). In Figure 6 all cases—humans, great apes, and fossils—have been assigned a dental age as if each was an ordinary white child from Ohio. For each subject, a single score (the coefficient of variation of human dental ages for the teeth I1-M2) measures concordance with ideal patterns in the reference population. An ideal average child would score zero—a score almost never achieved, even in the reference population itself—while highly discordant patterns should score high. Summary scores for humans and great apes are displayed as a histogram in Figure 6 (sample size is reduced because CV is calculated only for cases with six or seven teeth; a missing datum allowed for one incisor or one premolar only). Results are not surprising: Humans scored by human standards tend to lower, more concordant values than apes scored by human standards. Each distribution in Figure 6 has a long tail, but central tendencies that differ enormously. The median score is 10.9 for southern African black children and 30.6 for great apes (significantly different at $P < 0.0001$ by median test). The few fossils shown here (complete or near-complete data are required) appear to break into two groups: Neanderthal plus *Homo erectus* near the human median, and *Australopithecus* plus *Homo habilis* to the right, much nearer African apes, although in the area of overlap of the distributions.

The point of Figure 6 is simply this: There are many ways to show that patterns of dental development in humans, apes, and fossil hominids are not and were not all the same. Indeed, a world of comparisons becomes possible when dental development is quantified.

CONCLUSIONS

Tossing a coin, the simplest method of classifying individuals into two sets, gives the correct answer about half the time. Any reasonable method that adds information about central tendencies of the sets should better 50% in the long run, and better it substantially if the difference between the sets is large. In this light, the 92% error Lampl et al. (1993) attribute to Smith (1986) seems extraordinary. The present analysis of 665 human subjects scored by an outside investigator, however, does not sustain this attribution. The Smith (1986) method of classifying individual humans and African apes by statistical similarity to group standards of dental development (central tendency discrimination) showed reasonable success when tested on a pool of 659 *Homo sapiens,* 43 *Pan,* and 33 *Gorilla* juveniles. Ninety-two percent of humans and 88% of great apes were correctly discriminated from scores of dental development when both molar and incisor/canine fields were represented. The method declined in accuracy as sampling degraded, but the kind of teeth observed was more important than their number. As few as two teeth were adequate to distinguish more than 90% of humans and apes if molar and incisor-canine fields were sampled, whereas with two near or adjacent teeth, one might as well toss a coin.

Combinations of teeth that produce poor classification reflect real patterns of overlap between the dental development of humans and apes. Thus, data for the tooth pair M_2 and P_4, which produce very poor classification, were also shown to overlap extensively in a typical ape and a pool of humans. In contrast, anterior and posterior teeth, which occupy widely different development tracks in individual apes and humans, clearly distinguished the two.

Although few infant human subjects were available for study, the method was shown to be unreliable for 33 infant apes, where most cases were labelled "human." This is probably because great ape standards are substantially in error under 3 years of age, although special factors may also contribute towards error in this age group. The Dean and Wood (1981) growth standard for great apes needs improvement overall, but the present study suggests that the standard is much more accurate in pattern at older ages. Until these growth standards are improved, the SMITH86 program should not be used on infants.

Conflicts between the present study and that of Lampl et al. (1993) are not resolved by appeal to effects of population, age, or sample completeness. Limiting test cases to comparable sets in the two studies gives 52% success in Lamp et al. (1993) data (N = 23; Table 8) vs. 92% success in the present study (N = 531; Table 6). The probability that both results are random draws from the same universe is vanishingly small; it is much more likely that the two studies differ in aspects of methodology. It is clear that Lampl et al. (1993) used some unique protocols in assigning scores to specimens, but no complete resolution is possible without access to original specimens. In any case, replication can establish the success or failure of any method in the long run. To standardize methods for the future, central tendency discrimination of ape and human dental development (Smith, 1986) is now available as a computer program (to those who send a DOS-formatted diskette and a stamped, self-addressed mailer); others are encouraged to try it.

Lastly, a review of dental development of fossil hominids, restricting analysis to cases that satisfy criteria for adequate sampling, continues to find that dental development was not static during the evolutionary history of the Hominidae (Smith, 1986, 1991a, 1993, 1994). Australopithecines appear more distant from modern humans in patterns of dental development than *Homo erectus* and Neanderthals. Gracile australopithecines continue to be classified with apes by central tendency discrimination, although some other ways of looking at the problem suggest their dental development had shifted slightly in a human direction. The most compelling interpretation of shifts in developmental pattern continues to be that they reflect adaptation of overall rates of growth and aging (Schultz, 1960; Mann, 1975; Smith, 1992, 1994). Other lines of evidence concerning the dentition concur that human growth and aging evolved relatively recently (Bromage and Dean, 1985; Smith, 1989a). There is much to be learned by quantifying dental development; this study is only a beginning.

ACKNOWLEDGMENTS

I thank Robert Tompkins for lending data on human subjects to provide an independent test. Philip Gingerich wrote the computer program SMITH86. Kari Brandt assisted with data management and analysis. I thank Chris Dean for permission to study his collection of radiographs of great apes, and C.K. Brain, Chris Stringer, Richard Leakey, Elizabeth Vrba, Phillip Tobias, and the governments of Kenya and Tanzania for access to original fossil material in their care. I thank Alan Walker and Bernard Wood for casts and original radiographs, and Glenn Conroy, Philip Gingerich, and Bill Jungers for helpful comments on the manuscript. Tompkins' data were collected courtesy of C.B. Preston, M. Barter, Dr. Verwyen, A. Ferreira, E.T. Maganu, and the Ministry of Health of Botswana. This research was supported by National Science Foundation grants BNS 90-20974 and BNS 89-18011.

LITERATURE CITED

Anderson DL, Thompson GW, and Popovich F (1976) Age of attainment of mineralization stages of the permanent dentition. J. Forens. Sci. *21:*191–200.

Anemone RL, Watts ES, and Swindler DR (1991) Dental development of known-age chimpanzees, *Pan troglodytes* (Primates, Pongidae). Am. J. Phys. Anthropol. *86:*229–241.

Beynon AD, Dean MC, and Reid DJ (1991) Histological study on the chronology of the developing dentition in gorilla and orangutan. Am. J. Phys. Anthropol. *86:* 189–203.

Bromage TG (1987) The biological and chronological maturation of early hominids. J. Hum. Evol. *16:*257–272.

Bromage TG, and Dean MC (1985) Re-evaluation of the age at death of immature fossil hominids. Nature *317:* 525–527.

Broom R, and Robinston JT (1951) Eruption of the permanent teeth in the South African fossil apemen. Nature *167:*443.

Conroy GC (1988) Alleged synapomorphy of the M1/I1 eruption pattern in robust australopithecines and *Homo*: Evidence from high-resolution computerized tomography. Am. J. Phys. Anthropol. *75:*487–492.

Conroy GC, and Vannier MW (1987) Dental development of the Taung skull from computerized tomography. Nature *329:*625–627.

Conroy G, and Vannier MW (1991) Dental development in South African australopithecines. Part I: Problems of pattern and chronology. Am. J. Phys. Anthropol. *86:*121–136.

Dahlberg AA, Menegaz-Bock RM (1958) Emergence of the permanent teeth in Pima Indian children. J. Dent. Res. *37:*1123–1140.

Dean MC (1987) The dental developmental status of six East African juvenile fossil hominids. J. Hum. Evol. *16:*197–213.

Dean MC and Wood BA (1981) Developing pongid dentition and its use for ageing individual crania in comparative cross-sectional growth studies. Folia Primatol. *36:*111–127.

Demirjian A, Goldstein H, and Tanner JM (1973) A new system of dental age assessment. Hum. Biol. *45:*211–227.

Fanning EA (1962) Third molar emergence in Bostonians. Am. J. Phys. Anthropol. *20:*339–346.

Gremiatski MA, and Nesturkh MF (1949) Teshik Tash. Moscow: Moscow University (in Russian).

Grine FE (1987) On the eruption pattern of the permanent incisors and first permanent molars in *Paranthropus*. Am. J. Phys. Anthropol. *72:*353–359.

Kuykendall KL (1992) Dental Development in Chimpanzees *(Pan troglodytes)* and Implications for Dental Development Patterns in Fossil Hominids. Ph.D. dissertation, Washington University, St. Louis, MO.

Kuykendall KL, Mahoney CJ, and Conroy GC (1992) Probit and survival analysis of tooth emergence ages in a mixed-longitudinal sample of chimpanzees *(Pan troglodytes)*. Am. J. Phys. Anthropol. *89:*379–399.

Lampl M, Monge JM, and Mann AE (1993) Further observations on a method for estimating hominoid dental developmental patterns. Am. J. Phys. Anthropol. *90:*113–127.

Mann AE (1968) The Paleodemography of *Australopithecus*. Ph.D. dissertation, The University of California, Berkeley.

Mann A (1972) Hominid and cultural origins. Man *7:*379–386.

Mann AE (1975) Paleodemographic Aspects of the South African Australopithecinae. Philadelphia: University of Pennsylvania.

Moorrees CFA, Fanning EA, and Hunt EE (1963) Age variation of formation stages for ten permanent teeth. J. Dent. Res. *42:*1490–1502.

Schultz AH (1960) Age changes in primates and their modification in man. In JM Tanner (ed.): Human Growth. Oxford: Pergamon, pp. 1–20.

Simpson SW, Lovejoy CO, and Meindl RS (1990) Hominoid dental maturation. J. Hum. Evol. *19:*285–297.

Smith BH (1986) Dental development in *Australopithecus* and early *Homo*. Nature *323:*327–330.

Smith BH (1989a) Dental development as a measure of life history in primates. Evolution *43:*683–688.

Smith BH (1989b) Growth and development and its significance for early hominid behaviour. Ossa *14:*63–96.

Smith BH (1991a) Dental development and the evolution of life history in Hominidae. Am. J. Phys. Anthropol. *86:*157–174.

Smith BH (1991b) Standards of human tooth formation and dental age assessment. In MA Kelley and CS Larsen (eds.): Advances in Dental Anthropology. New York: Wiley-Liss, pp. 143–168.

Smith BH (1992) Life history and the evolution of human maturation. Evol. Anthropol. *1:*134–142.

Smith BH (1993) The physiological age of KNM-WT 15000. In A Walker and RE Leakey (eds.): The Nariokotome *Homo erectus* Skeleton. Cambridge, MA: Harvard University Press, pp. 195–220.

Smith BH (1994) Sequence of eruption of the teeth in *Macaca, Pan, Australopithecus,* and *Homo:* Its evolutionary significance. Am. J. Hum. Biol. *6:*61–76.

Tobias PV (1967) The Cranium of *Australopithecus (Zinjanthropus) boisei.* Olduvai Gorge, Volume 2. Cambridge, UK: Cambridge University Press.

Tompkins RL (1991) Relative Dental Development in Upper Pleistocene Fossil Hominids and Recent Humans. Ph.D. Dissertation, University of New Mexico, Albuquerque.

Virchow H (1920) Die Menschlichen Skeletreste aus dem Kämpfe'schen Bruch im Travertin von Ehringsdorf bei Weimar. Jena: Verlag von Gustav Fischer.

Weidenreich F (1936) The mandibles of *Sinanthropus pekinensis:* A comparative study. Palaeontol. Sin., series D, vol.7, fasc. 3.

Weidenreich F (1937) The dentition of *Sinanthropus pekinensis:* A comparative odontography of the hominids. Palaeontol. Sin., n.s. 1:120–180.

Wood BA (1991) Hominid Cranial Remains. Koobi Fora Research Project Volume 4. New York: Oxford University Press.