

METRICAL DISTINCTIVENESS OF INDIAN CRANIA DEMONSTRATED THROUGH STEPWISE MULTIVARIATE CLASSIFICATION

Pathmanathan Raghavan

School of Archaeology and Anthropology, College of Arts and Social Sciences, The Australian National University, Canberra 0200, Australia.

Email: ramapithecus@usa.net

Gayathiri Pathmanathan

Department of Anthropology, Panjab University, Chandigarh 160001, India.

Email: marshalwolf@yahoo.co.in

Keywords: South Asia, Indian craniometry, stepwise multivariate classification, index analysis

ABSTRACT

Aims and objectives. Application of discriminant function analysis to confirm the metrical distinctiveness between Indian crania and the cranial series measured by W.W. Howells from elsewhere across the world, including the use of stepwise analysis to investigate whether the indices that are diagnostic of Indian crania have implications for the efficiency and reliability of an Indian cranium's identification as Indian.

Materials and methods. The analyzed Indians include 1,002 adult crania, representing six language groups from northwest, North and South India, for which all 42 measurements used in the analysis are available. All measurements for the Indians and comparative Howells' series were converted to Mossiman indices to remove the effects of size. For the discriminant function analysis, the Indian crania were randomly divided into "classification" samples (for developing the discriminant formulae) and "prediction" samples (to which the discriminant formulae were applied). Measurements were entered stepwise into the analysis based on their overall discriminatory value.

Results and conclusion. Indian crania with diagnostically Indian index scores could be reliably identified as Indian with the entry of a smaller number of measurements than Indian crania whose index scores are atypical of Indians. However, with the entry of all 42 measurements, the great majority of Indian crania were

correctly identified regardless of their index scores. Accordingly, index scores have implications for the correct identification of Indian crania only when a limited number of measurements are available for analysis.

GOALS OF THIS PAPER

One goal of this paper is to demonstrate the value of craniometric index comparisons as part of using discriminant function analysis to detect the population affinities of human crania. When only a limited number of measurements is available for use, but their scope covers craniometric indexes that are indicative for the population to which the cranium belongs, then discriminant function analysis can be expected to detect the correct affinity to the degree that its index values align with those of its population. However, as long as critical measurements (covering the population's indicative indices) are available for analysis, then the larger the number of analyzed measurements, the more likely is detection of the correct affinity. These points are demonstrated through stepwise discriminant function analysis of crania of known population affinity from India. The results feed into the second goal of this paper, which is to show that a "South Asian" craniometric identity can be discerned at the level of the individual cranium, as well as the population level (as previously demonstrated by Raghavan *et al.* 2013).

INTRODUCTION

Raghavan *et al.* (2013) analyzed ten series from northern India and South India, based on measurements recorded by the first author. The northern Indian series comprised Punjabi and Haryanavi crania from Northwest India and Hindi crania from North India (all Indo-Aryan speakers), while the South Indian cranial series included Urdu, Konkani (Indo-Aryan speakers), Telugu, Kannada, Tamil, Tulu and Malayalam (Dravidian speakers). The series were measured using the linear chords and subtenses defined by Howells (1989). Averages for the males were compared on selected chords and illustrative indices with the series documented by Howells (1989) and with the Veddas of Sri Lanka (data from Woo and Morant 1934; Warusawithana-Kutilake 1996). Principal Components Analysis and Mahalanobis distance comparisons were also applied to the six Indian series with sample sizes (males and females combined) of more than 50.

The results demonstrate that northern Indians and South Indians are more similar to each other than either is to any of the Howells' series, which cover the inhabited world outside of South Asia. In terms of cranial shape, South Indians are particularly unlike any of the Howells' series, whereas northern Indians are intermediate between South Indians and "Caucasoid" series in Egypt and Europe. In terms of size, Indian crania are small, of a similar size to the crania of small-bodied populations such as Kalahari Bushmen and Andaman Islanders. Vedda crania are similar to Indian crania both in their small size and in their shape, as registered by the index comparisons, indicating that we may refer to a "South Asian" craniometric identity that includes both Indians and Veddas.

These results transcend the findings of previous studies where Indians are represented by generic samples and/or by a smaller selection of measurements. Brace *et al.* (1991) found that, in terms of cranial shape, the closest macropopulations to Indians are Europeans and American Indians, although a specific affinity with Andaman Islanders is evident if Europeans are excluded from analysis. In the cranial-shape analysis of Warusawithana-Kutilake (1996:Fig.

10) North Indians and Veddas are particularly close, and otherwise similar to Andaman Islanders, to the exclusion of Arabians and Burmese. Stock *et al.* (2007) found that their Ganges Valley, Indus Valley, Himalayan, Northeast India Tribal and South India/Deccan samples cluster tightly together on the basis of small cranial size and distinctive shape tendencies; the closest populations to Indians include Afghanistan/Iran, Andaman Islanders, Veddas, and to a lesser degree Sinhalese and West Asians. Similarly, in his analysis of our Punjabi sample, Wright (2008) found that the metrically closest population is Andaman Islanders, which themselves cluster with West Asians and Howells' Egyptian sample. An Indian-Andamanese affinity is a recurring finding of these studies, but not confirmed by Raghavan *et al.* (2013), which leaves it as an open question for this study.

An important aspect of the study by Raghavan *et al.* (2013) is the demonstration of Indians' craniometric variability. The six series subjected to Principal Components Analysis all include crania that are similar in size to the average recorded for European, African and southwest Pacific series, as well as other crania that are smaller than any documented for these latter series. Plots of the second and third components, the two main shape components, show at least some overlap between every Indian series and every Howells' series, with only a minority of Indian crania falling outside the range of variation of all of the Howells' series. Similarly, index analysis, extended to all ten Indian series, shows their considerable shape variability. With few exceptions, all of the Indian series include cranial vaults that are broad in relation to their length as well as very narrow in this respect, and cranial vaults that are low as well as very tall in relation to their length. All of the series also include faces that vary from very broad to narrow in relation to their height, with orbits that vary from broad to very tall and nasal apertures that vary from narrow to very broad. Comparable results were found in all of these aspects by Saini *et al.* (2017) in their study of adult Indian crania from two medical institutions (not covered by Raghavan *et al.* 2013) in the Hindi-speaking state of Uttar Pradesh.

Table 1 summarizes the average shape characteristics of Indian crania, compared to the Howells' series, as revealed through index analysis (Raghavan et al. 2013). Indian cranial vaults tend to be narrow (low cranial index) but, in the case of South Indians, also tall (high vault length-height index). The frontal bone tends to bulge (high frontal curvature index) while the parietal and occipital bones are intermediate in their curvature. Indians tend to be orthognathic (low gnathic index) and posteriorly broad in relation to facial breadth (high posterior cranio-

facial index). Indian faces tend to be narrow (high upper facial index), with weak projection of the zygomatic arches in relation to maxillary breadth (high bizygomatic-bimaxillary index), but intermediate in nasal aperture shape and, as a specifically South Indian feature, relatively broad in orbit shape (low orbital index). Finally, Indians' faces tend to be medially projecting, with high facial subtenses; the only facial flatness index for which Indians do not register a high average value is South Indians' maxillary flatness index, which is intermediate.

Table 1. Summary of Indians' index comparisons (after Raghavan et al. 2013). For explanation of acronyms, see Table 3.

Index	Northern Indians	South Indians
Cranial index (GOL:XCB)	Low	Low
Vault length-height index (GOL:BBH)	Intermediate	High
Frontal curvature index (FRC:FRS)	High	High
Parietal curvature index (PAC:PAS)	Intermediate	Intermediate
Occipital curvature index (OCC:OCS)	Intermediate	Intermediate
Gnathic index (BNL:BPL)	Low	Low
Posterior cranio-facial index (ZYB:ASB)	High	High
Transverse cranio-facial index (XCB:ZYB)	Intermediate	Intermediate
Upper facial index (ZYB:NPH)	High	High
Bizygomatic-bimaxillary index (ZYB:ZMB)	High	High
Nasal index (NLH:NLB)	Intermediate	Intermediate
Orbital index (OBB:OBH)	Intermediate	Low
Frontal (facial) flatness index (FMB:NAS)	Very high	Very high
Orbital flatness index (EKB:DKS)	Very high	Very high
Maxillary flatness index (ZMB:SSS)	High	Intermediate
Naso-dacryal index (DKB:NDS)	High	High
Simotic index (WNB:SIS)	High	High

The above discussion raises the question of whether Indian crania, despite their variability, can be consistently classified correctly through multivariate statistical methods. The average index tendencies described above have potential implications for the correct classification of Indian crania. For instance, Indian specimens with a very low cranial index may be clearly distinct from the great majority of the crania measured by Howells (with their predominantly intermediate to high cranial index), whereas Indian specimens with a high cranial index may be less readily distinguishable from non-Indian crania.

However, biometricians who undertake multivariate statistical analyses may regard shape as an aspect that can be appreciated only from considering the totality of measurements (e.g. Howells 1989:13–16), and accordingly dismiss the informative value of describing Indians' cranial shape tendencies through select index comparisons. Such a dismissal can be tested through the classification exercises undertaken in this paper. If it is valid, then there should be no implications from index considerations (Table 1) as to how well Indian crania can be correctly classified. If on the other hand we do find that index considerations have predictive value for how well Indian crania can be correctly classified, then we may infer that index comparisons indeed have value for describing shape tendencies.

MATERIALS AND METHODS

The six best-sampled Indian series measured by the first author, three from northern India and three from South India, are the series analyzed in this study. In this study, only the 1,002 specimens with all of their measurements intact are included for analysis, producing the sample sizes listed in Table 2. These crania are stored in various human anatomy and physical anthropology collections across India (Figure 1), along with three crania stored in the Museum of South Australia in Adelaide. The majority of the crania were obtained from medical dissections, with the remainder donated by collectors. All of these crania are of recent antiquity, which rules out the need to consider secular change in craniometric tendencies, which is an issue raised by Saini *et al.* (2017).

The great majority are from individuals of known adult status and sex, and in some cases their names are recorded. The first author also confirmed the adult status of the crania based on their degree of dental development and cranial suture closure, and confirmed their registered sex by recording the robustness of their supraorbital region, mastoid process and nuchal musculature with reference to the standard casts prepared by Larnach and Freedman (1964) for Australian Aboriginal crania. Thus, the first author familiarized himself with the adult male and adult female range of morphological variation of the various Indian series, allowing him to sex the adult crania of unrecorded sex.

Table 2. Sample sizes of the six Indian series included in the present study.

Language group	Location	Male sample size	Female sample size
Punjabis (Indo-Aryan)	Northwest India	94	49
Haryanavis (Indo-Aryan)	Northwest India	89	47
Hindis (Indo-Aryan)	North India	164	106
Telugu (Dravidian)	South India	62	46
Kannada (Dravidian)	South India	140	55
Tamils (Dravidian)	South India	96	54

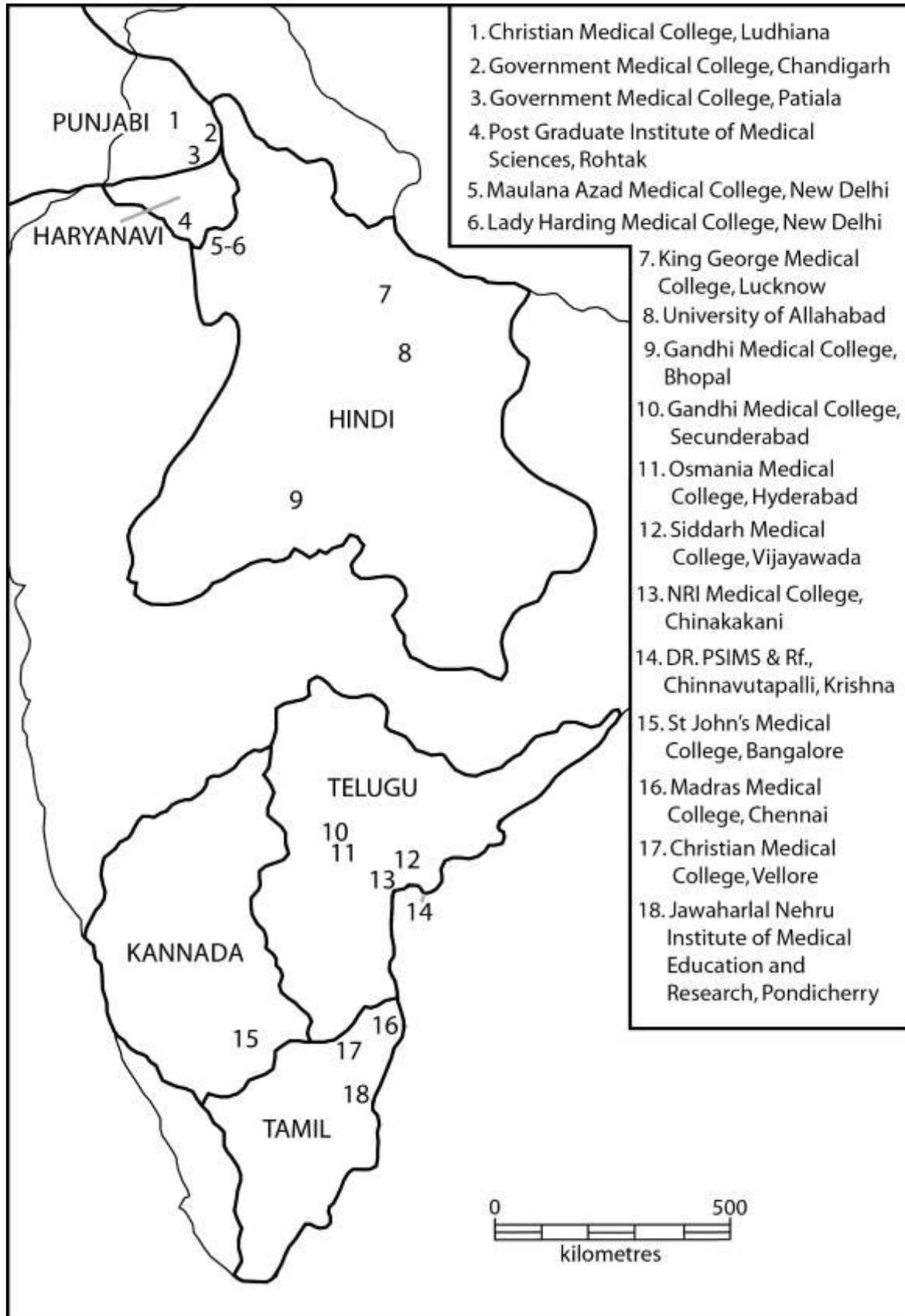


Figure 1. Distribution of the six language groups included in this study and location of the anatomical laboratories holding the analyzed skulls. Illustration by David Bulbeck.

The measurements included in the analysis are the 42 listed in Table 3, ordered according to their overall weights in the male analysis conducted here (which differ to some degree from their overall weights in the female analysis). These include the measurements used in the Principal Components Analysis of Raghavan *et al.* (2013), which also provides the means and standard deviations for the various Indian series. However, the current analysis excludes supra-orbital projection and glabella projection, as these measure zero on a small number of very gracile crania, thus preventing their inclusion in calculating the geometric mean of these crania's measurements. The geometric mean of the analyzed crania's measurements is used so as to adjust the measurements for overall size, specifically in the form of "Mosimann indices" (Darroch and Mosimann 1985) whereby each measurement is placed in the numerator and the specimen's geometric mean in the denominator. The effectiveness of this procedure in controlling for overall size can be seen from the fact that the overall weights of the measurements in the discriminant function analysis is unrelated to the scale of the measurement. For instance, the four measurements with the largest weights include three facial subtenses, which typically measure between 1 and 17 mm for any specimen, and also glabello-occipital length, which is usually a specimen's largest measurement (generally measuring between 160 and 190 mm).

The objective of discriminant function analysis is to assign weights to the entered measurements (here, Mosimann indices) that maximize the correct classification of the analyzed specimens to their population (cranial series). One potential objection to this technique is that it can be over-optimistic in its classification, in effect "overfitting" the available specimens based on the samples at hand. To counter this objection, and to take advantage of the large sample sizes of Indian crania (Table 2), the Indian samples were divided into "classification" and "prediction" groups. The average sample sizes for the series analysed by Howells (1989) is 48 for males and 44 for females. Accordingly, these

are the sample sizes allowed for the Indian "classification" groups, to mitigate against any tendency for the development of classificatory formulae that are biased in favor of the correct classification of what are sometimes very large samples of Indian crania. A random sample of 48 males and 44 females was selected from each of the Indian series for inclusion in their "classification" groups, and the remaining crania were assigned to their "prediction" groups.

Initial analysis included all of Howells' (2009) data for the male and female series in Howells (1989), along with the "classification" Indian crania, for the measurements listed in Table 3 (males and females analyzed separately). The default options in XLSTAT 2013 were applied including the option to calculate canonical variates. The "prediction" Indian crania were then classified to series based on the formulae developed for the "classification" Indian crania. When this exercise was performed allowing for differences between the covariance matrices, it resulted in perfect classification of all of the "classification" crania (Indian or non-Indian) but extremely poor classification of the "prediction" Indian crania (data not shown). In other words, the resulting formulae overfitted the specimens to their samples and thus lacked generality of application beyond these samples. Performing the same exercise with the simplifying assumption of equality of the compared series' covariance matrices, on the other hand, produced similar rates of correct classification for both the "classification" and the "prediction" Indian crania. Accordingly, the latter approach was followed as it neutralized the potential problem of overfitting.

The measurements differed in their contribution to the overall correctness of classification, as registered by the differences in their overall weights. Accordingly, stepwise classification was then undertaken, entering the measurements sequentially based on their weights in the overall analysis (the order shown in Table 3 for males, and a slightly different order for females). The classificatory success of the analysis was recorded at each step.

Table 3. Howells' measurements employed in the discriminant function analysis.

Measurement	Acronym
1. Simotic subtense (projection of the nasal bridge from the simotic chord)	SIS
2. Dacryon subtense (dacryon projection from biorbital breadth)	DKS
3. Naso-dacryal subtense (least nasal bone projection from interorbital breadth)	NDS
4. Maximum glabello-occipital cranial length	GOL
5. Orbital breadth from dacryon (left)	OBB
6. Simotic chord (least breadth across the nasal bones)	WNB
7. Maximum nasio-occipital cranial length	NOL
8. Nasal breadth	NLB
9. Bifrontal (upper facial) breadth	FMB
10. Interorbital breadth (across the dacrya)	DKB
11. Frontal subtense (maximum projection from the frontal chord)	FRS
12. Zygomaxillary subtense (subspinale projection from bimaxillary breadth)	SSS
13. Maximum transverse cranial breadth (above the supramastoid crests)	XCB
14. Malar subtense (maximum projection of malar bone from the maximum malar length)	MLS
15. Mastoid process breadth	MDB
16. Biauricular breadth (across the roots of the zygomatic processes)	AUB
17. Maximum transverse frontal breadth	XFB
18. Orbital height (left)	OBH
19. Bizygomatic facial breadth	ZYB
20. Inferior malar length (left)	IML
21. Bijugal breadth (breadth across the middle malars)	JUB
22. Cheek height (left)	WMH
23. Nasal height	NLH
24. Occipital subtense (maximum projection from the occipital chord)	OCS
25. Basion-nasion (cranial base) length	BNL
26. Biorbital breadth (across the ectoconchia)	EKB
27. Nasio-frontal subtense (nasion projection from bifrontal breadth)	NAS
28. Minimum cranial breadth (across the infratemporal crests)	WCB
29. Nasion-prosthion (upper facial) height	NPH
30. Foramen magnum (basion to opisthion) length	FOL
31. Parietal subtense (maximum projection from the parietal chord)	PAS
32. Bimaxillary (inferior malar) breadth	ZMB
33. Basion-bregma cranial height	BBH
34. Mastoid process height	MDH
35. Basion-prosthion (facial) length	BPL
36. Maximum malar length (left)	XML
37. Biasterionic (maximum occipital) breadth	ASB
38. Frontal (nasion to bregma) chord	FRC
39. Bistephanic breadth (frontal breadth across the inferior temporal lines)	STB
40. Occipital (lambda to opisthion) chord	OCC
41. Parietal (bregma to lambda) chord	PAC
42. External palate breadth	MAB

In addition, the implications of index considerations were incorporated into the stepwise classification based on the following protocol. Consider the case when the step involves the entry of the numerator measurement of an index for which northern and/or South Indians score relatively high in comparison with the Howells' series. Then the expectation is that any Indian specimens scoring particularly high on that index should have their correct classification assisted, whereas Indian specimens scoring particularly low on that index would be more prone to misclassification with a Howells' series. (The inverse expectations apply with the stepwise entry of the denominator measurement of an index for which Indians score low compared with the Howells' series.) Also, in the case of northern and/or South Indian crania that are outliers on an index for which Indians are intermediate, then the stepwise entry of either the index's numerator or denominator measurement (whichever comes first) should work against the correct classification of Indian crania. This is because it should tend to confuse the Indian specimen with the Howells' series that score either higher or lower than Indians for that index.

Application of the protocol proceeded as follows. The first step involved flagging the Indian crania more than two standard deviations above or below their series average with reference to the indices listed in Table 1. Then, to address those indices where northern and/or South Indian crania score high or low, inclusion of the index's numerator resulted in the following flags: a single plus for crania with high expectation of correct classification (for instance, being two standard deviations above their series average for an index where Indians scores high), or a double minus for crania with low expectation of correct classification. As for the indices where northern and/or South Indians are intermediate, then the inclusion of either the numerator or else the denominator measurement for the index resulted in any Indian specimen more than two standard deviations above or below its series average being flagged with a single minus. The plus and minus signs accumulated with the

stepwise inclusion of measurements into the analysis (Tables 4 and 5).

Some examples can help illustrate the procedure. At step 1, with the introduction of SIS, any Indian specimen with a particularly high simotic index (more than two standard deviations above its series average) would be flagged as "Indian+", while any Indian specimen with a particularly low simotic index (more than two standard deviations below its series average) would be flagged as "Indian--".

Let us now move onto step 2. Say the specimen is a northern Indian female and it is an outlier on its vault length-height index. Then, if it had been Indian+ at step 1 it would now be Indian+/- at step 2, or just "Indian" (with the plus and minus signs cancelling each other out), that is, not clearly expected to differ from the generality of specimens in its series in terms of its classificatory success. However, if the northern Indian female had been just "Indian" at step 1 it would now be Indian- at step 2, and if it had been Indian-- it would now be Indian---.

Say, instead that we are dealing with male specimens, and we are considering the implications of the introduction of DKS to the analysis at step 2. The range of possible flags for a male Indian are "Indian++" (particularly high on both their simotic and dacryon subtenses), "Indian+" (particularly high on just one of these two subtenses), "Indian" (within two standard deviations of its series average on both subtenses), "Indian-" (particularly high on one of the two subtenses and particularly low on the other), "Indian--" (particularly low on just one of the two subtenses) and "Indian---" (particularly low on both subtenses).

To simplify the analysis, at each step of the stepwise discriminant function analysis, any multiplicity of plus or minus signs is ignored. This produces three groups of Indian crania:

- "Indian+", with an expected higher than average classificatory success (the EHTACC group),
- "Indian", with an expected average classificatory success (the EACC group), and
- "Indian-", with an expected lower than average classificatory success (the ELTACC group).

Table 4. Implications of the indices for classifying northern and South Indian male crania.

Steps in the classification at which the indices are involved	Northern Indians		South Indians	
	High	Low	High	Low
Step 1: simotic subtense (SIS)	+	--	+	--
Step 2: dacryon subtense (DKS)	+	--	+	--
Step 3: naso-dacryal subtense (NDS)	+	--	+	--
Step 4: glabello-occipital length (GOL)	-	-	None	None
Step 5: orbital breadth (OBB)	-	-	None	None
Step 8: nasal breadth (NLB)	-	-	-	-
Step 11: frontal subtense (FRS)	+	--	+	--
Step 12: zygomaxillary subtense (SSS)	+	--	-	-
Step 13: maximum cranial breadth (XCB)	-, -- or ----	+, - or +/-	-, -- or ----	+, - or +/-
Step 18: orbital height (OBH)	None	None	--	+
Step 24: occipital subtense (OCS)	-	-	-	-
Step 27: nasio-frontal subtense (NAS)	+	--	+	--
Step 29: facial height (NPH)	+	--	+	--
Step 31: parietal subtense (PAS)	-	-	-	-
Step 32: bimaxillary breadth (ZMB)	+	--	+	--
Step 33: basion-bregma height (BBH)	None	None	+	--
Step 35: basion-prosthion length (BPL)	--	+	--	+
Step 37: biasterionic breadth (ASB)	+	--	+	--

Table 5. Implications of the indices for classifying northern and South Indian female crania.

Steps in the classification at which the indices are involved	Northern Indians		South Indians	
	High	Low	High	Low
Step 1: simotic subtense (SIS)	+	--	+	--
Step 2: glabello-occipital length (GOL)	-	-	None	None
Step 4: dacryon subtense (DKS)	+	--	+	--
Step 5: naso-dacryal subtense (NDS)	+	--	+	--
Step 7: frontal subtense (FRS)	+	--	+	--
Step 9: orbital breadth (OBB)	-	-	None	None
Step 15: maximum cranial breadth (XCB)	-, -- or ----	+, - or +/-	-, -- or ----	+, - or +/-
Step 16: zygomaxillary subtense (SSS)	+	--	-	-
Step 17: occipital subtense (OCS)	-	-	-	-
Step 18: nasal breadth (NLB)	-	-	-	-
Step 19: orbital height (OBH)	None	None	--	+
Step 20: nasio-frontal subtense (NAS)	+	--	+	--
Step 24: parietal subtense (PAS)	-	-	-	-
Step 26: bimaxillary breadth (ZMB)	+	--	+	--
Step 34: basion-bregma height (BBH)	None	None	+	--
Step 35: facial height (NPH)	+	--	+	--
Step 36: biasterionic breadth (ASB)	+	--	+	--
Step 38: basion-prosthion length (BPL)	--	+	--	+

The results of the stepwise classification are presented separately for males and females, to show that the two sexes produce similar results.

RESULTS

Classification results with all 42 measurements

The overall classifications of the Howells' series and Indian "classification" samples using all 42 measurements are presented in Table S1 for males and Table S2 for females. Overall correctness is 1,315/1,636 (80%) for males and 1,170/1,420 (82%) for females, with more than half of the crania from every series correctly classified and, in some cases, all of the crania correctly classified. Importantly, the "misclassifications" mostly involve populations that are known to be closely related, such as North Japanese classified as South Japanese and vice versa; or, on a broader scale, the "misclassifications" between the lowland East Asian males of Anyang, Hainan, the Philippines, North and South Japan, and Taiwan (Atayal). However, there are also occasional cases of gross misclassification, such as a Berg (Switzerland) male classified as a Kalahari Bushman.

The Indian samples differ from the Howells' series in their overall lower correct classification (192/288 or 67% of males, 167/264 or 63% of females). This largely reflects the frequent cross-classifications between the three northern Indian samples, which are very similar to each other (239/276 or 87% of northern Indians classified with a northern Indian series), and between the three South Indian samples, which also are very similar to each other (241/276 or 87% of South Indians classified with a South Indian series). There is also some degree of cross-classification between the northern Indian and South Indian samples, consistent with the finding of Raghavan *et al.* (2013) that the northern and South Indian series comprise a distinct Indian craniometric cluster. The Kannada are always classified with an Indian series, but the other Indian series include small numbers of specimens classified with one or the other non-Indian series in East Asia, the Pacific, Africa and Europe. Finally, there is just one case of a

non-Indian specimen being classified as Indian, namely a Norse female classified as Hindi.

The classification results of the "prediction" Indian crania (Tables S3 and S4) provide a realistic view of the metrical distinctiveness of Indian cranial series, because the classificatory formulae were developed for the "classification" Indian crania. The first point to note is that, as long as the sample size is 11 or more, the modal classification is always the correct classification. The second point to note is that the clear majority are classified to their correct region in India (74% of northern Indian crania are classified with a northern Indian series including 31/51 Punjabis, 141/178 Hindis and 30/44 Haryanavis, while 82% of South Indian crania are classified with a South Indian series including 12/16 Telugu, 51/58 Tamils and 83/103 Kannada). Finally, 409/450 or 91% of these Indian crania overall are classified with an Indian rather than a non-Indian series (233 of 273 or 85% of the northern Indian crania, and 176 of 177 or 99% of the South Indian crania). Interestingly, of the 41 cases classified with a non-Indian series, 21 (51%) were classified with a European or Egyptian series.

Stepwise classification results

The stepwise classification results are presented in combination with the index analysis. As explained previously, based on index analysis the Indian crania were sorted into three groups with, respectively, an expected higher than average correct classification (EHTACC), expected average correct classification (EACC), and expected lower than average correct classification (ELTACC).

The point of stepwise classification is to find the minimum set of measurements that achieves classificatory results that are essentially as correct as the full measurement suite. This minimum set of measurements is marked by reaching a plateau along which the inclusion of additional measurements has minimal impact on classificatory correctness. The practical significance of these plateaus is illustrated well with the results for the Indian "classification" crania.

For the EHTACC females, a stable result of 100 per cent correct classification was achieved

at step 21 with the introduction of FMB (Figure 2). As for the EACC females, 96–98 per cent correct classification became stable at step 28 with the entry of FOL. Classificatory correctness of the ELTACC females stabilized at step 29 (with the entry of MDH), and at a lower rate (about 91 per cent correct classification). Thus, the index analysis not only accurately distinguished between female Indian crania with higher than average, average, and lower than average correct classification, it also flagged classificatory efficiency (in the sense of requiring fewer measurements entered stepwise into the analysis to achieve stable classificatory correctness).

For the EHTACC males, a stable result of approximately 100 per cent correct classification was achieved at step 13 with the introduction of XCB. The most notable departure from this ideal result occurred at step 38 (with the entry of FRC), at which point the correct classification of the EHTACC males dipped to 97 per cent (Figure 3). As for the EACC males, the plateau in classificatory results was delayed till step 19 (with the entry of ZYB) and the classificatory success was generally lower (about 95–98 per cent). As for the ELTACC males, the plateau in classificatory results was still further delayed (step 29, with the introduction of NPH) associated with relatively low classificatory correctness (about 92 per cent). Thus, index analysis accurately distinguished the Indian crania with expected lower than average correct classification from their counterparts with average or higher than average correct classification. Also, index analysis flagged classificatory efficiency in terms of the number of measurements (least for the EHTACC males, and most for the ELTACC males) required for stepwise entry into the analysis to achieve stable classificatory correctness.

Turning to the “prediction” Indian females, we can see that an excellent result was obtained for the EHTACC crania (Figure 4). A stable outcome of 100 per cent perfect classificatory correctness was achieved at step 13 with the entry of WNB. On the other hand, there was no clear difference between the EACC and ELTACC Indians in this analysis. Both groups

plateaued at around 86 per cent classificatory correctness at step 21 with the entry of FMB.

In the case of the “prediction” Indian males, the results were less even than was the case with the classification Indian males (Figure 5). On the one hand, the EACC Indian males achieved a plateau of about 92–94 per cent correctly classified at step 20, with the entry of IML. Further, a lower correct classification, of about 85 per cent, became stable at a later step (32, involving ZMB) for the ELTACC Indian males. However, for the EHTACC Indian males, there was no sign of stability in the level of classificatory correctness until step 35 (BPL), and then at about the same level (94 per cent) as the EACC males.

It is also of interest to review the proportions of Indian crania that are correctly classified to their region, that is, northern Indians classified with a northern Indian series and South Indians with a South Indian series. The predictions of index analysis in terms of expected higher than average, average and lower than average classificatory correctness are not as strong as for their being classified with an Indian series, because northern and South Indians have more similarities than differences in terms of their index analysis implications (Tables 4 and 5).

With the classification Indian females (Figure 6), classification to the same region accorded with the predictions of the index analysis at every step as of the introduction of OCS at step 17. However, to the degree that a plateau in classificatory correctness can be discerned, it occurred at a later step. These were step 23 (IML, about 93 per cent correctness) for the EHTACC females, step 29 (MDH, about 83 per cent correctness) for the EACC females, and step 30 (BNL, about 76–79 per cent correctness) for the ELTACC females.

With the classification Indian males (Figure 7), considering firstly the EACC crania, we see that a consistent result of 90–95 per cent classificatory correctness was achieved at step 27 with the entry of NAS. At step 28 (NAS) and the following steps, a similar outcome was obtained for the EHTACC crania, but with some up and down movement proceeding from step to step. Both groups were clearly distinguished from the ELTACC crania, whose classificatory

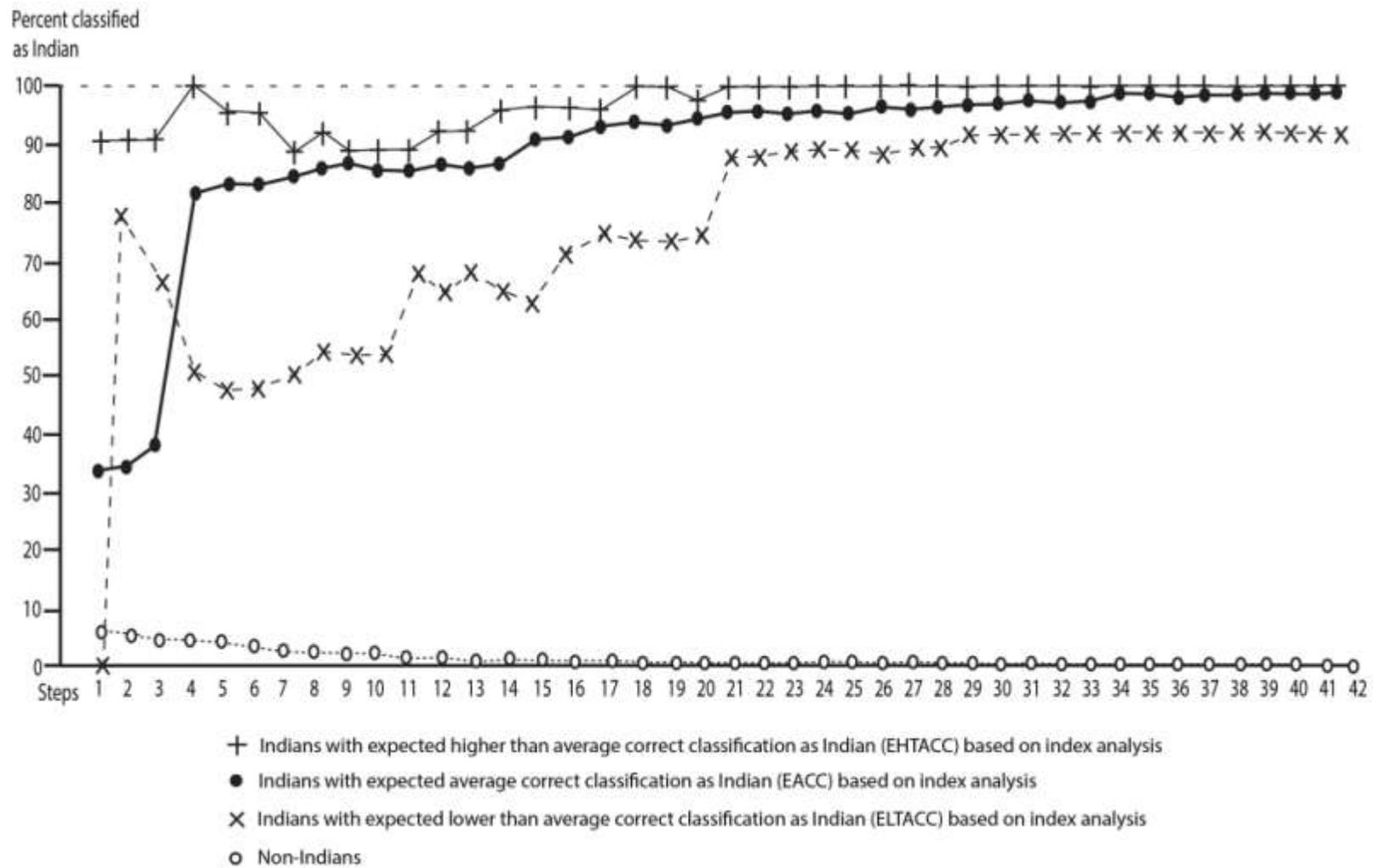


Figure 2. Female "classification" crania: proportions classified as Indian (stepwise classification). Illustration by David Bulbeck.

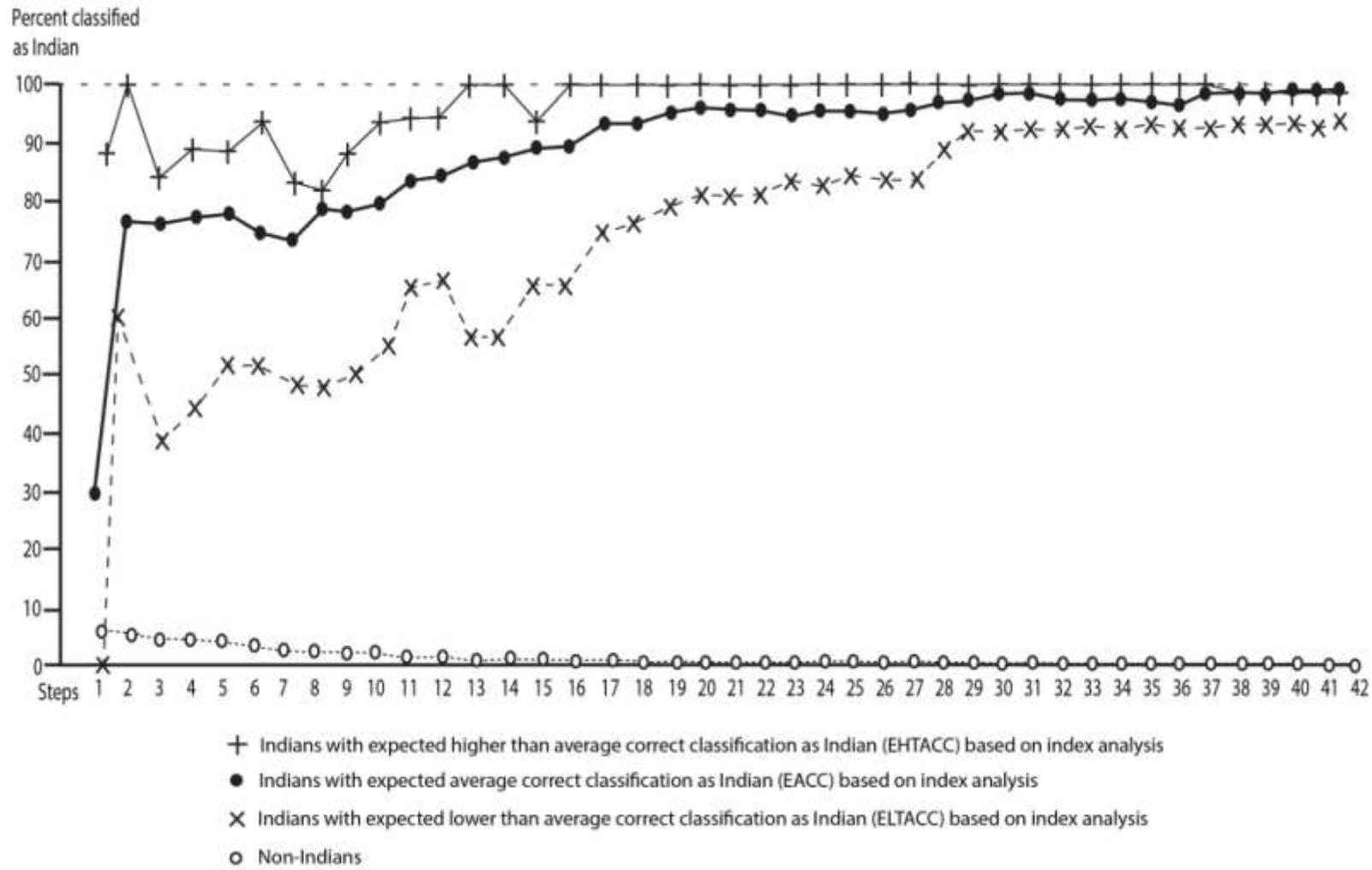


Figure 3. Male "classification" crania: proportions classified as Indian (stepwise classification). Illustration by David Bulbeck.

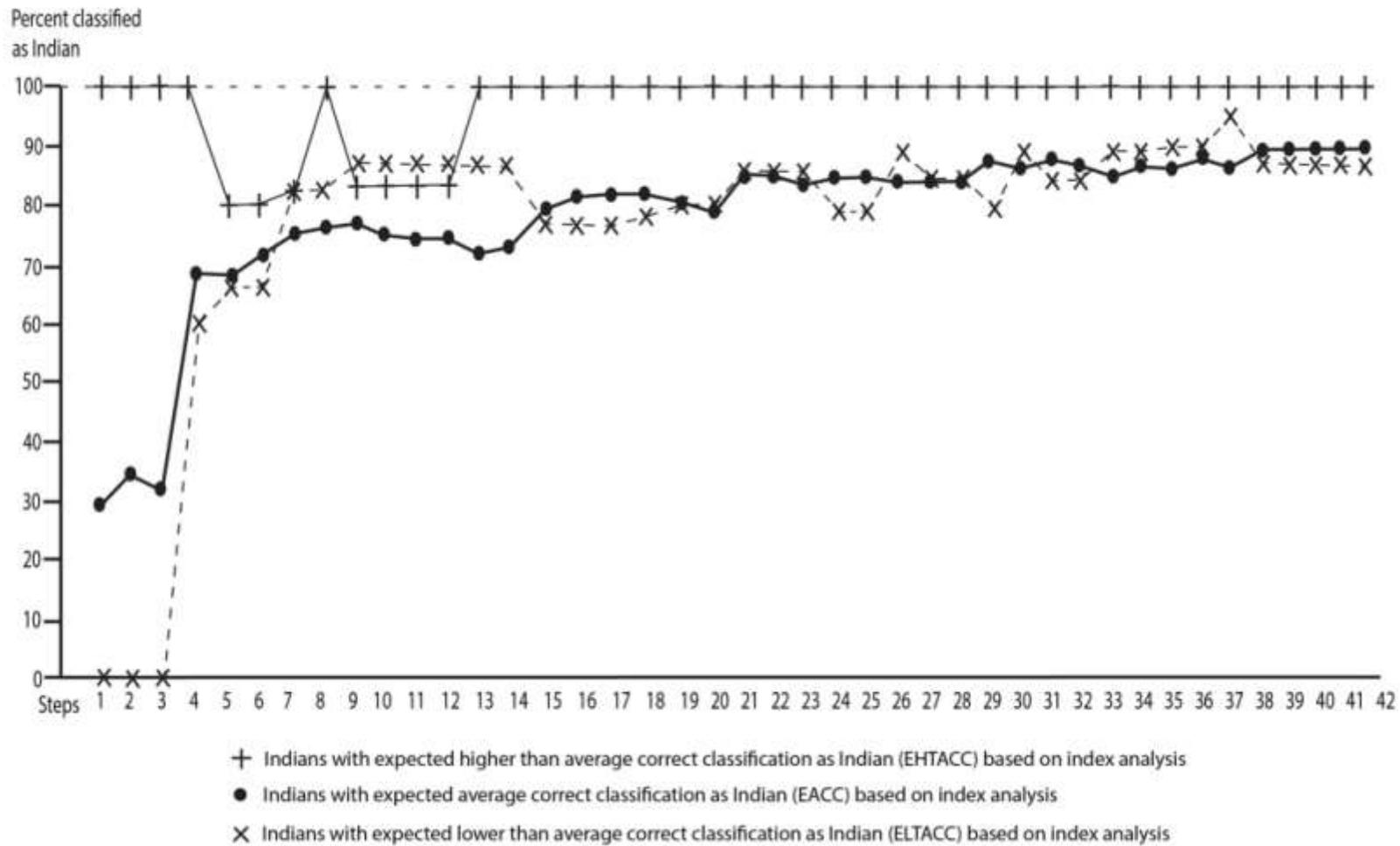


Figure 4. Female "prediction" Indian crania: proportions classified as Indian (stepwise classification). Illustration by David Bulbeck.

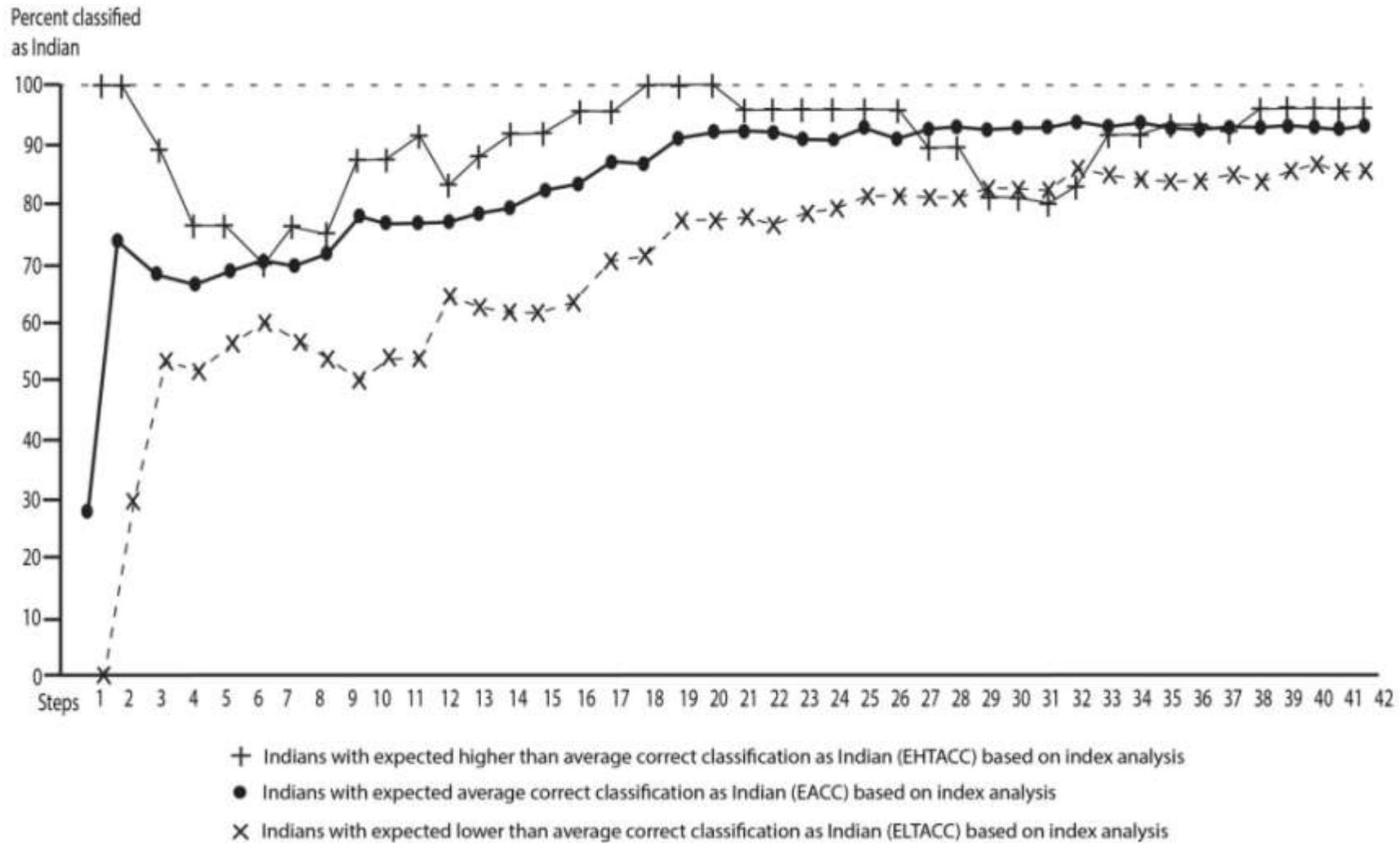


Figure 5. Male "classification" crania: proportions classified as Indian (stepwise classification). Illustration by David Bulbeck.

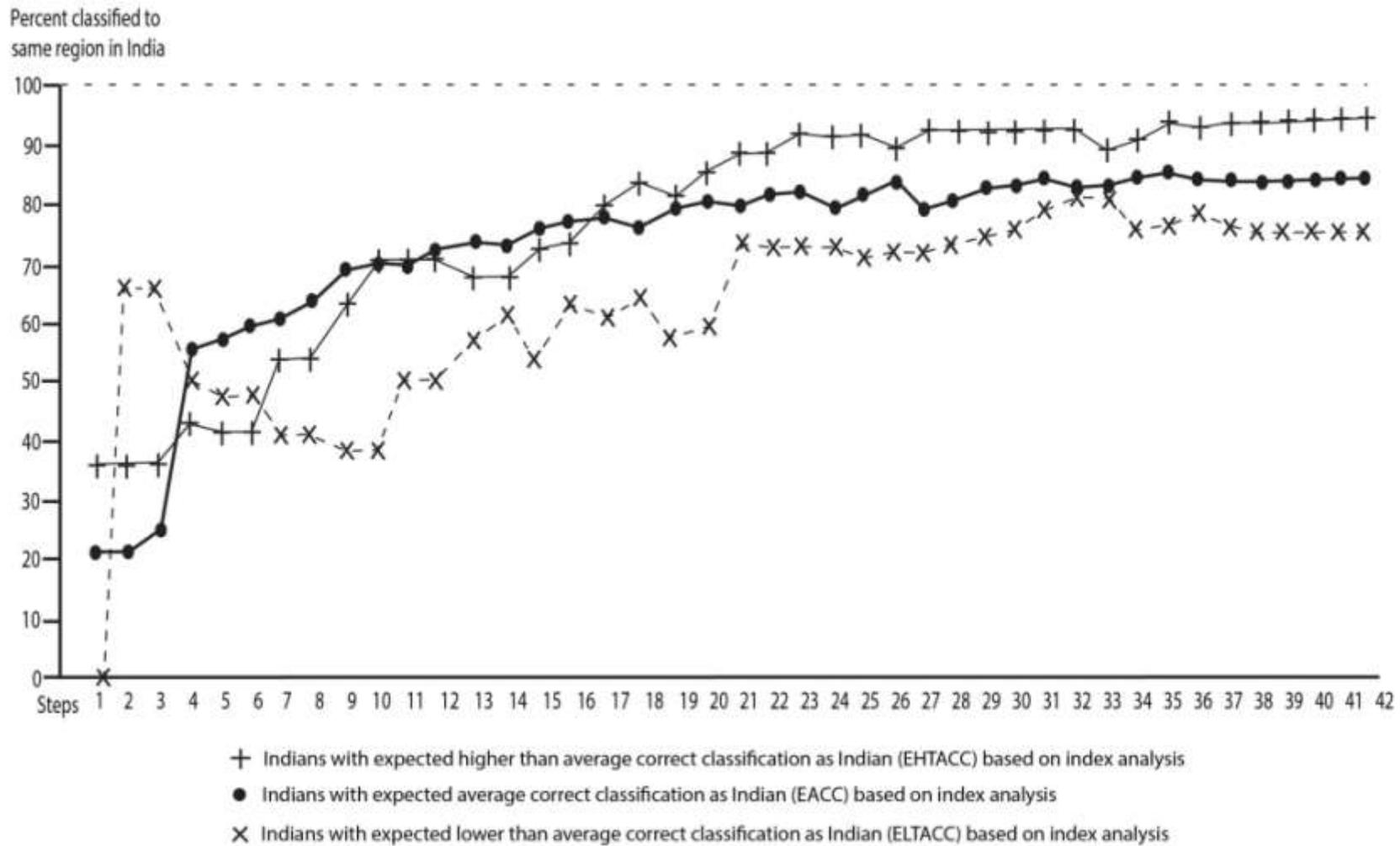
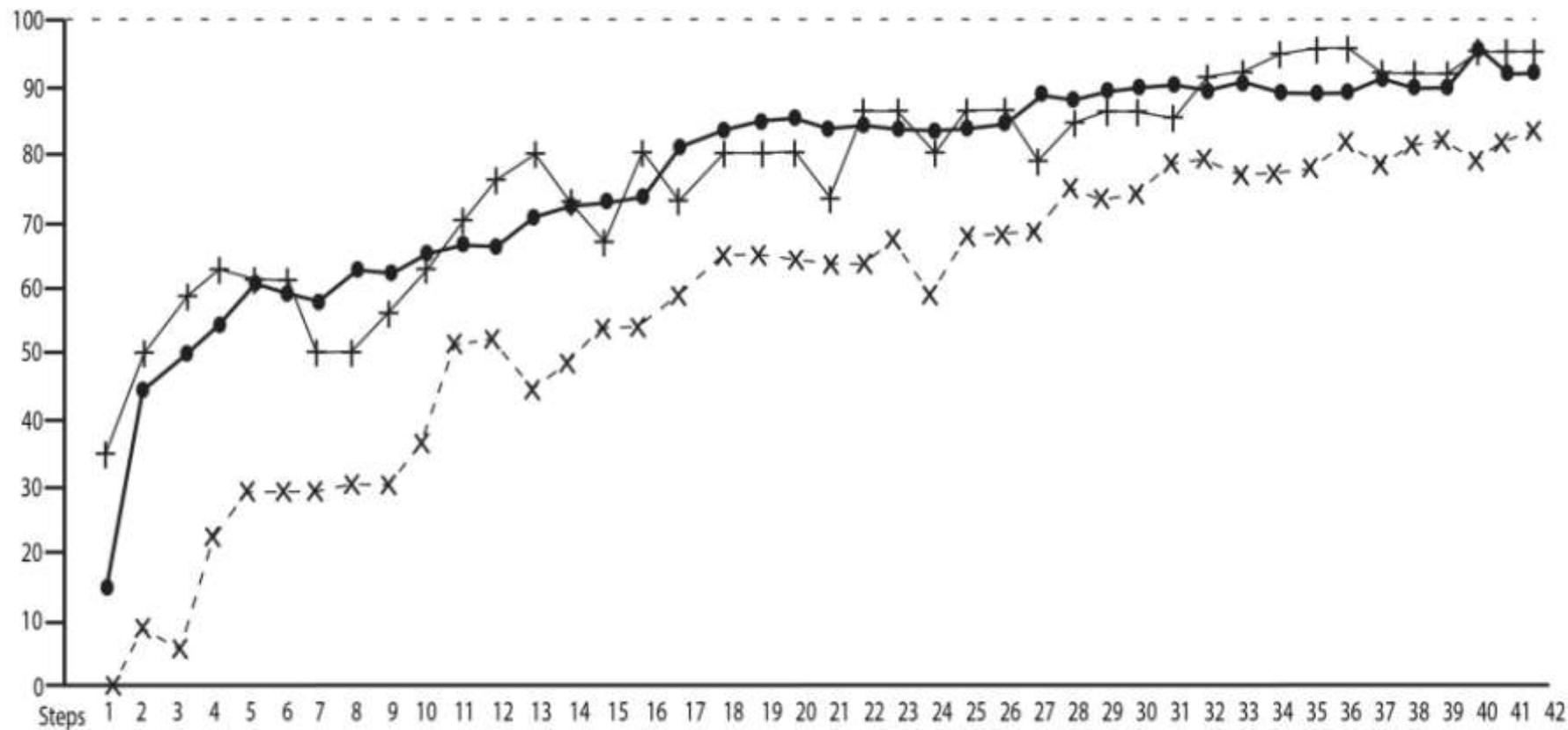


Figure 6. Female “classification” Indian crania: proportions classified to same region in India (stepwise classification). Illustration by David Bulbeck.

Percent classified to same region in India



- + Indians with expected higher than average correct classification as Indian (EHTACC) based on index analysis
- Indians with expected average correct classification as Indian (EACC) based on index analysis
- × Indians with expected lower than average correct classification as Indian (ELTACC) based on index analysis

Figure 7. Male “classification” Indian crania: proportions classified to same region in India (stepwise classification). Illustration by David Bulbeck.

Percent classified to same region in India

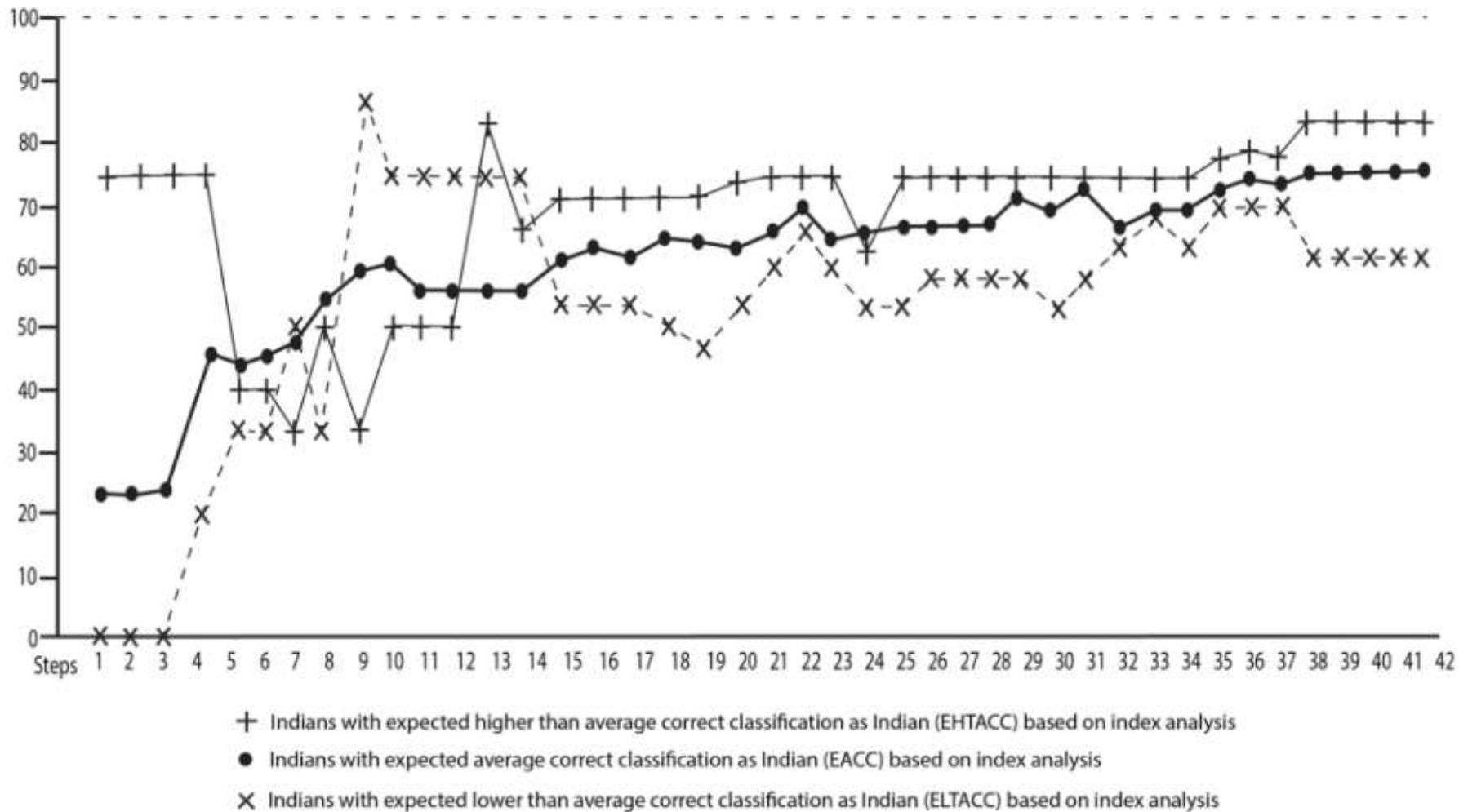


Figure 8. Female "prediction" Indian crania: proportions classified to same region in India (stepwise classification). Illustration by David Bulbeck.

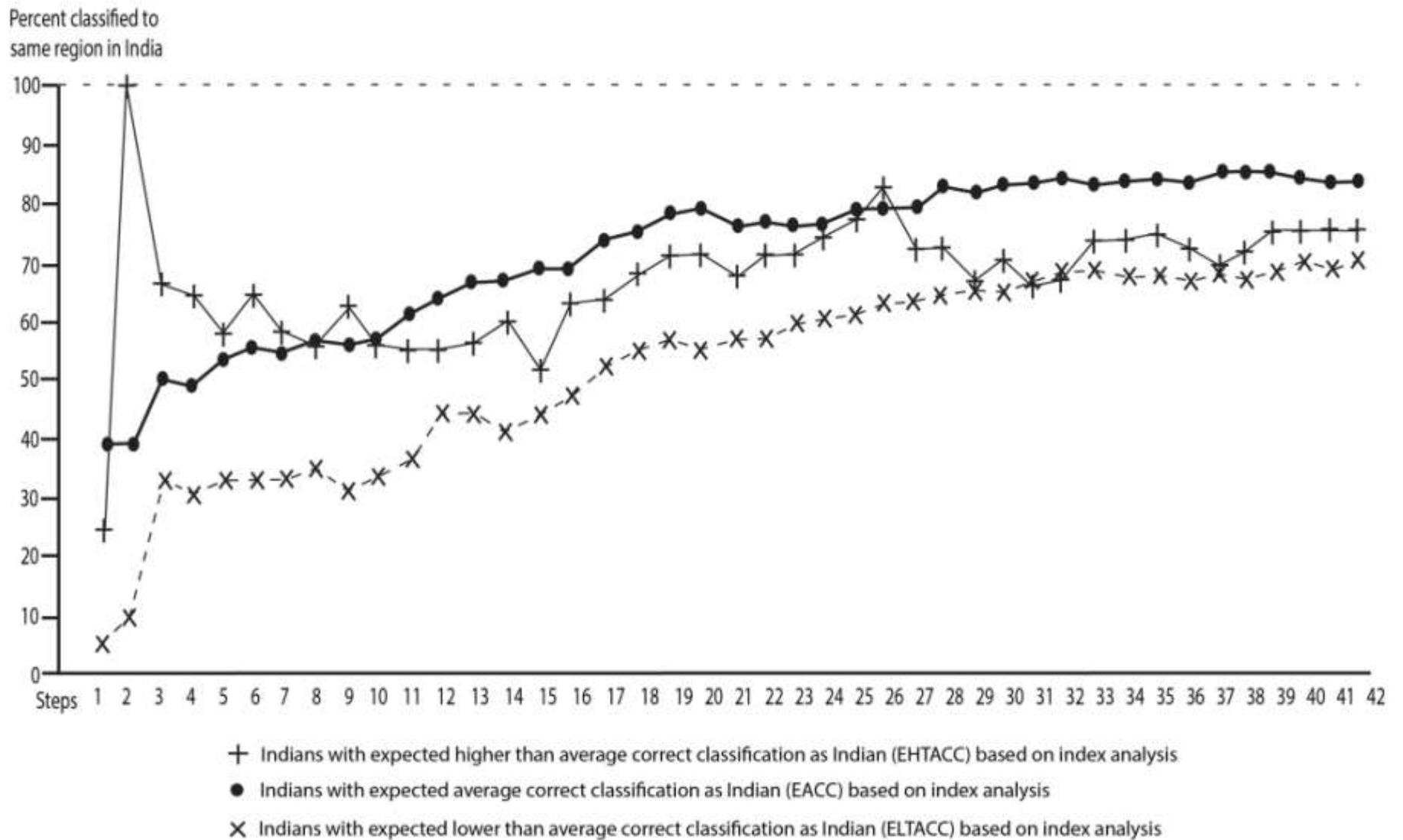


Figure 9. Male “prediction” Indian crania: proportions classified to same region in India (stepwise classification). Illustration by David Bulbeck.

correctness plateaued at about 75 per cent at step 28 (involving the entry of WCB).

As for the prediction Indian females (Figure 8), at most of the steps the proportion of correctly classified EHTACC crania was higher than the proportion of correctly classified EACC crania, which in turn was higher than the proportion of correctly classified ELTACC crania. However, it was only at step 26 (with the entry of ZMB) that these differentiations stabilized, and only at step 38 (with the introduction of BPL) that plateaus of classificatory correctness (85 per cent for the EHTACC females, 75 per cent for the EACC females and 61 per cent for the ELTACC females) were arrived at.

As for the prediction Indian males (Figure 9), it was the EACC group that was most successfully classified to the correct region. A stable outcome of around 83 per cent correctness was achieved at step 28 with the introduction of WCB. As for the ELTACC crania, a stable outcome was obtained at a later step (step 31, with the entry of PAS) and at a lower level (about 68 per cent correctness). Any stability in the proportion of correctly classified EHTACC crania is not in evidence until step 33 (BBH), after which point the proportion of correctly classified crania (around 74 per cent) fell between the proportions obtained for the EACC and ELTACC crania.

Finally, we may note how the stepwise analysis demonstrates the non-Indian status of the non-Indian crania. As the non-Indian crania measured by Howells were all treated here as “classification” crania, presentation of their classificatory results is restricted to Figures 2 and 3. The proportion of non-Indian crania misclassified with an Indian series dropped to about 5% at step 3 and virtually 0% at step 18. Overall, when non-Indian crania were misclassified, it was overwhelmingly with other non-Indian series rather than with an Indian series (Tables S1 and S2), underlining the distinctiveness of the Indian “craniometric identity”.

DISCUSSION

Implications for modern humans’ craniometric classification

Correct classification of the Indian series was merely moderate, involving 65 per cent of the “classification” crania and 45 per cent of the “prediction” crania. This can be ascribed to internal variability within the Indian series which made it difficult to develop classificatory formulae that are distinctive for each of the Indian series. One cause for this internal variability could be the wide geographic spread of the sampled Indian series and the anatomical institutes containing their crania, particularly affecting the Hindi and Telugu samples (Figure 1). By contrast, most of the Howells’ (1989) series involved specimens from a single cemetery. However, it should be noted that all of the Haryanavi crania came from a single institute, yet they did not differ markedly from the other Indian series in their classificatory correctness (Tables S1 and S2). Additional contributors to the internal variability of the Indian series, such as their being represented by multiple castes and their possibly heterogeneous origins (Majumder and Basu 2015), may also be at stake.

The classificatory analysis found a heuristic benefit in recognizing northern and South Indian groupings, with 87 per cent of the classification crania and 77 per cent of the prediction crania classified to their correct region. However, this would not justify treating northern India and South India as two homogeneous regions in terms of human biology. Both regions include large numbers of “tribal” societies represented by, at best, very small cranial samples in Indian anatomical institutes, which led to the exclusion of these societies from the study by Raghavan *et al.* (2013). Further, four of the groups reviewed in that study, including two South Indian groups that speak Indo-Aryan languages, were represented by samples of inadequate size for the multivariate analysis undertaken here.

The distinctiveness of the Indian series from the series measured by Howells (1989) may partly reflect the geographical distance from India of the Howells’ series. Of these, the closest are Andaman Islanders to the east and Cauca-

soid groups to the west. Andaman Islanders and Caucasoids were also the groups that previous craniometric studies generally found to be close to South Asians, as reviewed in our Introduction.

However, our classification results provide no confirmation for inferring a craniometric similarity between Indians and Andamanese. Of the 1,002 analyzed Indian crania, only two (both Punjabi “prediction” males) were classified as Andamanese. Similarly, none of the 64 Andamanese crania measured by Howells were classified with Indian groups (Tables S1 and S2). Projection of the mid-facial skeleton is a strongly distinguishing feature of South Asians (Tables 4 and 5), in contrast with the much flatter faces of Andamanese (Raghavan *et al.* 2013). Inclusion of facial subtenses in our analysis but to a lesser degree, or not at all, in previous studies can explain why our analysis disputes an Andamanese-South Asian craniometric affinity. Where the present study enhances the findings of Raghavan *et al.* (2013) is in showing that the difference between Andamanese and South Asians applies at the level of individual crania as well as the population level. Although Reich *et al.* (2009) found a deeply rooted genetic affinity between Andaman Islanders and South Indians, our evidence suggests that a specialized, distinct craniometric shape evolved for both Andamanese and South Indians as a result of genetic isolation between these populations in the distant past.

On the other hand, some level of craniometric similarity between Indians and Caucasoids is suggested by the fact that around half (51 per cent) of the misclassifications of the prediction Indian crania involved a Caucasoid series. Thus, the Eastern Mediterranean region lying between South Asia and Europe may be expected to include populations intermediate in their craniometrics between Indians and Caucasoids, as also found by previous studies, notably (Wright 2008). This finding aligns with the evidence from human genetics of an external contribution to the gene pool of Indo-European speakers distributed from Europe to northern India, associated with Holocene dispersal from Central Asia (Majumder and Basu 2015; Silva *et al.* 2017).

A deeper insight into Indians’ craniometric distinctiveness is provided by stepwise classification of the crania, aligned with analysis of the index implications (Figures 2 and 3). In overall terms, index analysis was successful in predicting which Indian crania would have, respectively, higher than average, average, and lower than average classificatory correctness—especially when using just a small number of measurements relating to indices that are particularly diagnostic of Indians. Further, the non-Indian status of the great majority of crania from the Howells’ series becomes quickly apparent with the entry of even a small number of stepwise measurements.

On the other hand, a small number of index values unusual for Indians may be a poor indicator that the cranium in question is non-Indian. As long as a sufficiently large suite of measurements was included in the analysis, correct classification rates of circa 80–90 per cent were obtained even for Indian crania with certain very non-Indian characteristics (Figures 2 to 5). Accordingly, multivariate analysis extracts far more information relevant to correctly classifying Indian crania than would be obtained from simply considering the index implications summarized in Table 1.

Our results have implications beyond the correct classification of Indian crania. First, any series that are craniometrically distinct should be amenable to index analysis. For instance, 284/297 (93 per cent) of Southwest Pacific (Australian, Tasmanian and Tolai) crania are correctly classified to a Southwest Pacific series (Tables S1 and S2), so there should be a combination of indices that reflects a distinctly Southwest Pacific cranial shape. Secondly, if our task were to assess the ancestry of a “mystery” cranium, then we should apply discriminant function analysis to the largest available suite of measurements, whilst exercising care in interpreting the results depending on which measurements are available. For instance, the question of a possible Southwest Pacific ancestry for the mystery cranium would be addressed by the available measurements only to the degree that they have a promising diagnostic

prognosis based on analysis of distinctively Southwest Pacific indices.

Implications for Indians' biological relationships

The implications of this study for the biological relationships of Indians to other branches of *Homo sapiens* are essentially the same as those discussed by Raghavan *et al.* (2013). The craniometric distinctiveness of Dravidian-speaking South Indians from every group except northern Indians is clearly confirmed. Northern Indians for their part show some similarity with “Caucasoid” (circum-Mediterranean) populations, on top of their strong similarities with South Indians. Accordingly, northern India is transitional between circum-Mediterranean and South Indian populations in their cranial shape, although with their major orientation southwards, towards the rest of the South Asian subcontinent. The craniometric distinctiveness of Indians accords with genetic studies that point to a predominantly indigenous component in Indians' ancestry, tempered with genetic influx from the north, associated in particular with the spread of Indo-Aryan languages (Reich *et al.* 2009; Silva *et al.* 2017).

CONCLUSION

Previous metrical study of Indian crania by Raghavan *et al.* (2013) from across the South Asian subcontinent found great variability within all of the series, combined with average trends whereby the Indian series resemble each other and contrast with series outside of South Asia. This contribution confirms the craniometric distinctiveness of Indian crania, notwithstanding their intra-series variability, through stepwise discriminant function analysis of the six with the largest sample sizes. The proportion of crania from places beyond South Asia classified as Indian is negligible, while over 90% of Indian crania are correctly identified as Indian. This correct identification of Indian crania is enhanced amongst those specimens whose possession of pronouncedly Indian craniometric attributes is revealed through index analysis. Indians' craniometric distinctiveness aligns with genetic evidence for the predomi-

nantly indigenous ancestry of Indians who speak Indo-Aryan and especially Dravidian languages.

ACKNOWLEDGEMENTS

Funding for the first author's fieldwork was provided by a large Australian Research Council grant to David Bulbeck and Colin Groves (“The Contribution of South Asia to the Peopling of Australasia”) and funding to the first author from the Australia-India Council in Canberra. We would also like to thank the many researchers from various institutions who assisted the first author with access to the collections, and especially the curators of the collections under their care, as detailed in Raghavan *et al.* (2013). Adam Lauer and an anonymous referee assisted the improvement of this article with their recommended amendments.

REFERENCES

- Brace, C.L., D.P. Tracer and K.D. Hunt. 1991. Human craniofacial form and the evidence for the peopling of the Pacific. *Indo-Pacific Prehistory Association Bulletin* 11:247–269.
- Darroch J.N. and J.E. Mosimann. 1985. Canonical and principal components of shape. *Biometrika* 72:241–252.
- Howells, W.W. 1989. *Skull Shapes and the Map: Craniometric Analysis in the Dispersion of modern Homo*. Cambridge, MA: Harvard University.
- Howells, W.W., 2009. The William W. Howells craniometric data set. Available at <http://web.utk.edu/~auerbach/HOWL.htm>
- Larnach, S. L., and L. Freedman. 1964. Sex determination of Aboriginal crania from coastal New South Wales. *Records of the Australian Museum* 26:295–308.
- Majumder P.P. and A. Basu. 2015. A genomic view of the peopling and population structure of India. *Cold Spring Harbor Perspectives in Biology*. Available at <https://cshperspectives.cshlp.org/content/7/4/a008540.full>
- Raghavan, P., D. Bulbeck, G. Pathmanathan and S.K.Rathee. 2013. Indian craniometric variability and affinities. *International Journal*

- of *Evolutionary Biology*. Article ID 836738. Available at <http://www.hindawi.com/journals/ijeb/2013/836738/>.
- Reich, D., K. Thangarej, N. Patterson, A.L. Price and L. Singh. 2009. Reconstructing Indian population history. *Nature* 471:489–494.
- Saini, V., M. Mehta, R. Saini, S.N. Shamal, T.B. Singh and S.K. Tripathi. 2017. Is north Indian population changing in craniofacial form? A study of secular trends in craniometric indices and its relation to sex and ancestry estimation. *Forensic Science and Criminology* 2(2):1–14, doi: 10.15761/FSC.1000115.
- Silva, M., M. Oliveira, D. Vieira, A. Brandão, T. Rito, J.B. Pereira, R.M. Fraser, B. Hudson, F. Gandini, C. Edwards, M. Pala, J. Koch, J.F. Wilson, L. Pereira, M.B. Richards and P. Soares. 2017. A genetic chronology for the Indian subcontinent points to heavily sex-biased dispersals. *BMC Evolutionary Biology* 17:88, doi: 10.1186/s12862-017-0936-9.
- Stock, J.T., M.M. Lahr and S. Kulatilake. 2007. Cranial diversity in South Asia relative to modern human dispersals and global patterns of human variation. In: M.D. Petraglia and B. Allchin (ed.), *The Evolution and History of Human Populations in South Asia*, pp. 245–268. New York, NY: Springer.
- Warusawithana-Kutilake, S. 1996. Cranial Variation and the Dispersal of Modern Humans in South Asia. Colombo: Tharansee Prints.
- Woo, T.L. and G.M. Morant. 1934. A biometric study of the “flatness” of the facial skeleton in man. *Biometrika* 26:196–250.
- Wright, R. 2008. Detection of likely ancestry using CRANID. In: M. Oxenham (ed.), *Forensic Approaches to Death, Disaster and Abuse*, pp. 111–122. Bowen Hills, Queensland: Australian Academic Press.