JOURNAL OF CONTROVERSIAL IDEAS

Article

Could 19th-Century Authors Have Noticed Bergmann’s “Rule” in Humans?

Robert Asher [1,*]

[1] Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK
* Corresponding author: r.asher@zoo.cam.ac.uk

Submitted: 10 January 2024, accepted: 2 October 2024, published: 30 October 2024

Abstract: The correlation of increasing size with latitude, known as “Bergmann’s rule”, was first articulated in the 1840s, but its potential applicability to humans was not recognized for another century. In this paper, I have tested if human craniometric data collected by 19th-century naturalists supported this “rule”. At least in the northern hemisphere, they did. Bergmann recognized a relationship between size and latitude in the 1840s, but others studying humans did not, possibly because they were preoccupied with applying anatomical data to debates about human intelligence. Links between cranial anatomy and racist dogma have long been debunked and profound similarities across human populations show that ethnic prejudice has no basis in evolutionary biology. Nonetheless, human populations are not homogeneous or less subject to evolutionary processes than other organisms. Some of these processes are evident in the datasets collected by 19th-century naturalists, whatever their socio-political views may have been.

Keywords: Cranial volume; *Homo sapiens*; phenotypic variation; latitude; Friedrich Tiedemann; Samuel George Morton; Stephen Jay Gould

How to cite: Asher, R.J. Could 19th-Century Authors Have Noticed Bergmann’s “Rule” in Humans? *Journal of Controversial Ideas* **2024**, *4*(2), 18; doi:10.35995/jci04020018.

Introduction

A number of geographical phenomena consistently influence phenotype in many animal species. Latitude and elevation directly impact temperature and thereby have obvious adaptive relevance. Like any other species, *Homo sapiens* has an ideal range of habitable temperatures. To permit survival and reproduction outside of that range, humans have many behavioral and phenotypic responses. As in other vertebrates (Lindsey 1966; Ashton et al. 2000; Meiri and Dayan 2003; Freckleton et al. 2003), humans exhibit latitude-related variation in size, which tends to vary proportionally with distance from the equator. Specifically, cranial size (Beals et al. 1984; Relethford 2004, 2010; Foster and Collard 2013; Katz et al. 2016) tracks overall body size (Gustafsson and Lindenfors

2009) and tends to be larger in populations from cold climates. Such populations typically also have a more rounded cranial shape (Beals 1972; Beals et al. 1983; Hubbe et al. 2009). While many factors affect size, a well-documented influence across vertebrate species is thermal efficiency. Structures retain heat better when they are large and round, minimizing the ratio of surface area to mass, compared to when they are small and angular (Beals 1972; Beals et al. 1984; Ocobock 2023).

Such latitudinal variation in animals is often presented as one or more "rules" (Meiri 2011; Stulp and Barrett 2016). According to "Bergmann's rule", body size increases at high latitudes with correspondingly colder temperatures. Bergmann's 1847 article in *Göttinger Studien* (the 1848 reprint of which is cited here) was the first to articulate a heat-retention mechanism behind this pattern. Numerous other factors influence body size besides adaptations for thermoregulation, including founder effects (Mori et al. 2019), neutral processes such as drift (Relethford 2010; Roseman 2016), selection for sexual dimorphism (Gigliotti et al. 2020), nutrition, phylogeny, or insular evolution, and make deviations from such "rules" fairly common (Geist 1987; Watt et al. 2010; Meiri 2011; Stulp and Barrett 2016; Bogin et al. 2022). Inferring cause and effect is therefore tricky, and a more appropriate descriptor might be "pattern" rather than "rule" (Meiri 2011). Whatever the causal mechanisms may be, changes in body size and shape at ever higher latitudes are reasonably well-documented in many animal species (Lindsey 1966; Ashton et al. 2000; Meiri and Dayan 2003; Meiri 2011; He et al. 2023), including humans (Beals 1972; Beals et al. 1984; Relethford 2004, 2010; Gustafsson and Lindenfors 2009; Foster and Collard 2013; Katz et al. 2016; Pomeroy et al. 2021; Ocobock 2023).

Bergmann (1848) formulated his observations on latitudinal size variation not long after Whewell (1837) coined the term "scientist". Studies of human craniometric variation from around the same time (e.g., Tiedemann 1837; Morton 1849) might in principle also have noticed correlations between size and latitude, yet they did not. There were some hints from naturalists such as Buffon and Blumenbach that climate played a role behind human variation (Gould 1996), but past studies of human morphology were focused on how they might indicate some aspect of intelligence or sociocultural value (Richards 2018), as opposed to investigating discrete, measurable attributes. Some used human morphological data to disprove (Tiedemann 1837), others to justify (Morton 1849), widespread racial stereotypes. These authors lacked concepts of evolution, genetics, heritability, and a statistical framework to evaluate their data, and each had his own historical context in which he collected and evaluated data. In hindsight, it is not surprising that early naturalists did not fully grasp many factors behind the data they observed. However, writing with similar conceptual resources, Bergmann (1848: 55–56) did articulate what remains a widely observed pattern: "It may not be as clear as we might wish, and irregularities do occur, but it is nonetheless obvious, that [from his perspective in the northern hemisphere] generally the larger species are farther north and the smaller farther south". (Appendix 1 provides the original German for this and other quotes.)

The question addressed in this study

By focusing across multiple species and not just humans, Bergmann (1848) had a large body of comparative data with which to formulate his ideas on the relationship of size to temperature and latitude. The data published by Tiedemann (1837) and Morton (1849) are limited to humans, but nonetheless comprised individuals from over 60° north and 40° south. In theory, they might also have asked if the patterns noticed by Bergmann were evident in their own data. However, unlike Bergmann, Tiedemann and Morton were preoccupied with debates about human racial superiority. Indeed, recent authors (e.g., Gould 1978, 1996; Michael 1988; Kaplan et al. 2015; Weisberg and Paul 2016; Mitchell 2018; Mitchell and Michael 2019; Bogin et al. 2022) have also been preoccupied by the social context of human data. To their credit, Lewis et al. (2011), in their critique of Gould (1978, 1996), acknowledged previous work showing a relation between cranial capacity and latitude (Beals et al. 1984), as did Lieberman (2001). However, they still focused on the implications of racism and investigator bias surrounding interpretations of human craniometric data. Whether or not the data published by Tiedemann (1837) and Morton (1849) reflect patterns of latitudinal variation in humans remains unknown. Here, I attempt to fill this gap by asking if the data collected by Tiedemann (1837) and Morton (1849) contained a signal supportive of Bergmann's pattern. Manica et al. (2007) sampled a much larger dataset, in turn derived from Hanihara and Ishida (2001, 2005). For comparison, I ask if it too reflects Bergmann's pattern.

Questions not addressed in this study

Asking if 19th-century investigators could have noticed a pattern in their data does not require re-hashing the copious literature (e.g., Beals et al. 1984; Ashton et al. 2000; Meiri and Dayan 2003; Freckleton et al. 2003; Relethford 2004, 2010; Foster and Collard 2013; Katz et al. 2016; He et al. 2023; Ocobock 2023) on the extent to which cold temperatures, typical of high latitudes, are causally linked to evolutionary increase in size. Nor does it require the assumption that climate is the only or even primary influence. There are many potential factors behind phenotypic change (e.g., Relethford 2010; Mori et al. 2019; Gigliotti et al. 2020 and others as cited above) and most species will probably be subject to multiple such factors over time.

Some previous investigations of human cranial (e.g., Roseman 2016) and postcranial (Savell et al. 2022) form have emphasized the importance of population structure in interpreting correlations between latitude and phenotype. These authors identified some skeletal variables that were significantly correlated with a climatic gradient, but still emphasized the role of neutral evolution behind variation in the modern human skeleton. A study focusing on postcranial indices of stature (Pomeroy et al. 2021) also emphasized neutral changes inherent in population history, but still supported a significant relation between climate and body proportions, in particular that environmental productivity (their "net primary productivity" [NPP]) and "temperature are generally the most consistent predictors" of the anthropometric indices they examined (Pomeroy et al. 2021: 11). Katz et al. (2016) focused specifically on human cranial variation, used methods to account for population history, and similarly concluded that climate had a significant affect on phenotype: "cold climate groups tend toward larger crania ... [Our] results are consistent with earlier descriptions of increased brachycephaly (higher cephalic indices) in cold-climate groups (Beals 1972). ... [B]oth size and shape changes are associated with climate differences" (Katz et al. 2016: 597).

While there are qualifications, recent studies focusing on explicit datasets sampling both human and non-human species remain generally consistent with Beals et al. (1984), and ultimately Bergmann (1848), in support of an association between climate and size. Moreover, a role for thermal efficiency does not contradict the observation that "evolution of the human cranium is complex and no one evolutionary process has prevailed at the expense of all others" (Roseman 2016: 582). Here, I do not attempt to reevaluate the causal basis of Bergmann's pattern, or discern the relative importance of the many factors that could have played a role in patterning the data collected by Tiedemann (1837) and Morton (1849), if indeed there are any non-random patterns at all. Instead, I address the latter question: is there a non-random signal present in the data collected by these

19th-century authors which, in turn, could have led them to notice clinal variation in humans?

The importance of early 19th-century data

If contemporary methods show no association of size with latitude in their data, then neither Morton nor Tiedemann would have had an evidential basis to consider factors such as those articulated by Bergmann, or look beyond a presumed relation between cranial capacity and intelligence. Alternatively, if their data do exhibit a size–latitude correlation, this would suggest that biological patterns are still discoverable in datasets collected nearly 200 years ago, despite the very different social and political standards of the naturalists who assembled them. At least in Morton's case, this would contradict the assertion of Kaplan et al. (2015: 22) who "challenge[d] the premise ... that Morton's confused data can be used to draw any meaningful conclusions." Historical data have intrinsic value, and the assessment of such data enables current generations to better appreciate how our predecessors understood humanity's place in nature.

Methods

Bergmann (1848) lacked major concepts of statistics and evolution, and intuited a relationship between climate and body size without quantitative techniques. He discerned a north–south cline in body size in part because he considered his data using geographic criteria. Accordingly, I quantified data from Tiedemann (1837) and Morton (1849) using latitudes derived from the localities they reported for each skull (see Appendix 2). All skulls from the same degree latitude, rounded to the nearest integer, were grouped into a single datapoint regardless of their longitude or reported ethnicity. For example, this joined skulls from Quebec City, North Dakota, Sault St Marie, La Roche-sur-Yon, and Zürich from the Tiedemann dataset into a mean for 47° north, and skulls from Mexico and Mumbai from the Morton dataset into a mean for 19° north. Doing so combines genetically distinct human populations (which as a species already exhibit low levels of genetic diversity; see Osada 2015) into single datapoints oriented around the focal question of this study: does latitude correlate with size? It thereby reduces what would otherwise be an inflated estimate of the degrees of freedom, were individual skulls to be treated as if they were independent data points. Pooling in this fashion reduced the number of samples (Appendix 3) from 514 to 56 in Morton and from 366 to 60 in Tiedemann.

Bergmann lacked an understanding of population history and data independence, but he did have an intuitive appreciation of ranks, which enabled him to make the connection between size and latitude. Simply put, Bergmann noticed when large specimens of a given species tended to occur at high latitudes. Non-parametric tests such as Kendall's *Tau* are appropriate to test the significance of ranked variables such as size and latitude and do not assume a normal distribution. Other approaches to test if 19th-century datasets contained a biological signal include conversion of continuous variables into categorical ones in the form of a chi-squared test, and examining co-variation in distance matrices in the form of a Mantel test. Both techniques can help determine if the frequency of large vs. small skulls in high vs. low latitudes deviates from random.

Results

Patterns in Tiedemann's and Morton's data

Cranial capacities published by Tiedemann (1837) and Morton (1849) increase along with latitude in the northern but not southern hemisphere (Figure 1A; Table 1), a result similar to that reported by Foster and Collard (2013). Indeed, Morton's data show the reverse trend in the southern hemisphere, albeit with just one skull over 40° south and 14 between 30° and 40° south. The positive correlation of cranial capacity and latitude in the northern hemisphere using global data from both datasets is significant using non-parametric, rank-based (Table 1) and Mantel (Table 2) tests. There is no correlation between latitude and size for the Tiedemann data in the southern hemisphere, and a significant, inverse correlation for Morton's data using either his own measurements or re-measurements of the same skulls by Lewis et al. (2011). Re-measurements from Lewis et al. (2011) recover a similarly significant, positive correlation for Morton's data in the northern hemisphere (Table 1).

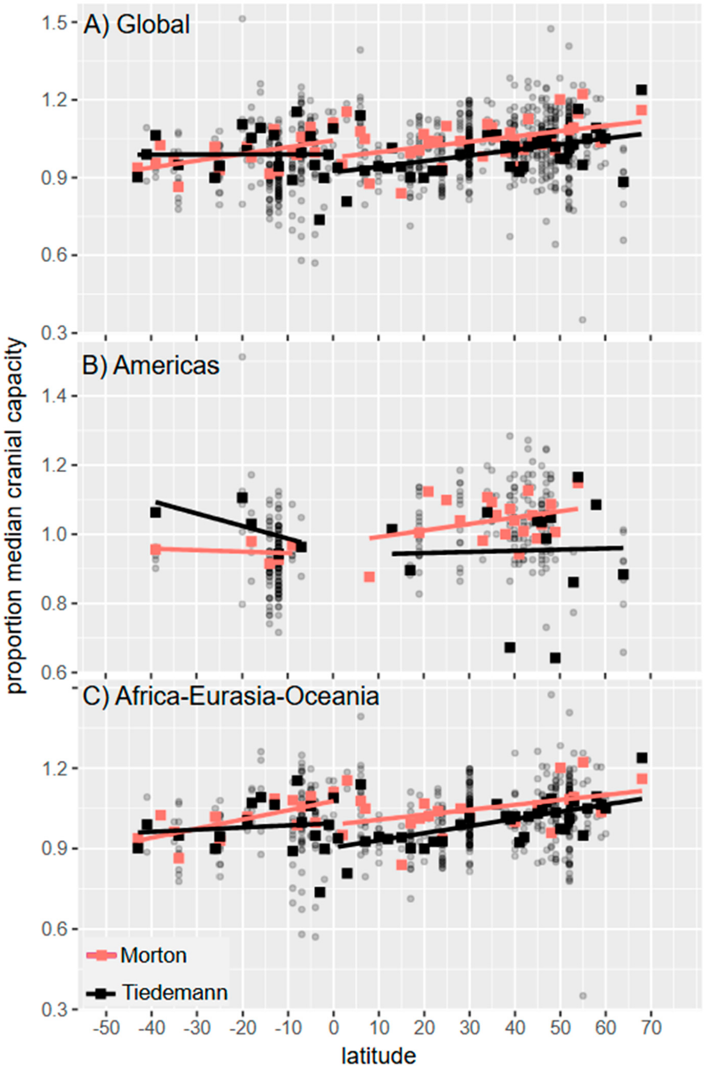

Figure 1. Cranial volumes expressed as a proportion of the median (*y*-axis) calculated separately for Morton (red/light shading) and Tiedemann (black) datasets globally (**A**), in the Americas (**B**), and rest of world (**C**). Lines show best fit linear models for northern and southern hemispheres, derived from squares representing mean cranial capacities per degree latitude rounded to the nearest integer (*x*-axis). Circles in background represent individual skulls.

Table 1. Correlations of cranial volume with latitude or climate, given as Kendall's rank correlation *tau* and Pearson product-moment coefficients. "Correlate" shows if comparison was between size and distance from equator ("latitude" averaged per degree), mean temperature ("temp"), or seasonality ("season"). "Conf int" refers to 95% confidence intervals for Pearson's *R*. "Morton-Lewis" indicates re-measurements in cm^3 of Morton's skulls by Lewis et al. (2011). "AM" = Americas. "RW" = rest of world. N = sample size. *P* values show significance of Kendall's rank correlations at far right, in bold when below alpha of 0.05.

data	correlate	N	tau (*R*, 95% conf int)	*p* for Tau
Tiedemann north	latitude	39	0.371 (0.494, 0.21 to 0.7)	**0.0009**
Tiedemann south	latitude	21	−0.01 (0.003, −0.43 to 0.43)	0.988
Morton north	latitude	38	0.265 (0.437, 0.14 to 0.66)	**0.0194**
Morton south	latitude	18	0.354 (0.533, 0.09 to 0.8)	**0.0407**
Morton-Lewis north	latitude	35	0.297 (0.396, 0.07 to 0.64)	**0.0117**
Morton-Lewis south	latitude	18	0.451 (0.608, 0.2 to 0.83)	**0.0085**
Tiedemann	temp	60	−0.238 (−0.327, −0.54 to −0.08)	**0.0075**
Tiedemann	season	60	0.23 (0.289, 0.04 to 0.51)	**0.0101**
Morton	temp	56	−0.115 (−0.233, −0.47 to 0.03)	0.2158
Morton	season	56	0.114 (0.211, −0.05 to 0.45)	0.2209
Morton-Lewis	temp	53	−0.091 (−0.204, −0.45 to 0.07)	0.3373
Morton-Lewis	season	53	0.072 (0.19, −0.09 to 0.43)	0.4514
Manica north	latitude	50	0.512 (0.639, 0.44 to 0.78)	**1.565×10^{-7}**
Manica south	latitude	18	−0.137 (−0.396, −0.73 to 0.08)	0.4543
Tiedemann north RW	latitude	35	0.469 (0.615, 0.35 to 0.79)	**7.4×10^{-5}**
Tiedemann south RW	latitude	18	0.059 (0.11, −0.38 to 0.54)	0.7332
Tiedemann north AM	latitude	13	0.513 (0.32, −0.53 to 0.57)	0.8072
Tiedemann south AM	latitude	5	−0.6 (−0.65, −0.97 to 0.54)	0.1416
Morton north RW	latitude	22	0.278 (0.417, −0.006 to 0.713)	0.071
Morton south RW	latitude	14	0.442 (0.688, 0.25 to 0.89)	**0.0283**
Morton north AM	latitude	22	0.056 (0.32, −0.12 to 0.65)	0.7139
Morton south AM	latitude	5	0 (−0.19, −0.92 to 0.83)	1

Table 2. Results from a Mantel test for correlation between two distance matrices as implemented with "mantel" in the vegan package for R (Legendre and Legendre 2012; Oksanen et al. 2022). For each dataset (i.e., Tiedemann 1837; Morton 1849; Manica et al. 2007), I constructed two matrices, one for distances in size and another for distances in latitude (rounded to the nearest degree). I took the mean of all skulls at each degree latitude, and tested northern and southern (plus equator) hemispheres separately. Ranks in the two matrices from lowest to highest distance significantly co-vary when the test statistic is at an extreme of a distribution based on (in this case) 999 iterations that randomly alter elements in one of the matrices (Sokal and Rohlf 1995: 816). I inferred a correlation coefficient (Mantel's *R*) using Kendall's rank sums. Significance is based at an alpha level of 0.05 (in bold) and corresponds to the number of randomized iterations with test statistics beyond that derived from the non-randomized distance matrices.

author	hemisphere	Mantel's *R*	Significance
Tiedemann	north	0.2035	**0.001**
Morton	north	0.1097	**0.016**
Manica	north	0.1944	**0.001**
Tiedemann	south	−0.7651	0.938
Morton	south	0.05916	0.154
Manica	south	0.1578	**0.02**

Decomposing the sample further into the two longitudinally widest, habitable regions (the Americas and Africa-Europe) revealed two departures from the pooled samples. First, the Tiedemann sample in North America (13° to 64°) showed no trend in cranial capacities with increasing latitudes (Figure 1B). Second, Tiedemann showed an upward, non-significant trend for cranial capacities with increasing latitude in South America (-7° to -39°; Figure 1B) when considered separately from the downward trend evident in the rest of the world (Figure 1C; Table 1).

I also tested if cranial volume correlates with seasonality and mean annual temperatures. Again, individual skulls were averaged so that each degree of latitude contributed one mean cranial capacity, one mean annual temperature, and one index of seasonality. The Tiedemann dataset shows significant correlations of cranial size with both; i.e., cranial capacity increases as seasonal changes in temperature increase, and as annual mean temperatures decrease (Figure 2; Table 1). The Morton dataset shows the same trends, but with datapoints averaged for each degree of latitude, and using either Morton's own measurements or re measurements from Lewis et al. (2011), neither seasonality nor mean temperature is significantly correlated with cranial capacity (Table 1).

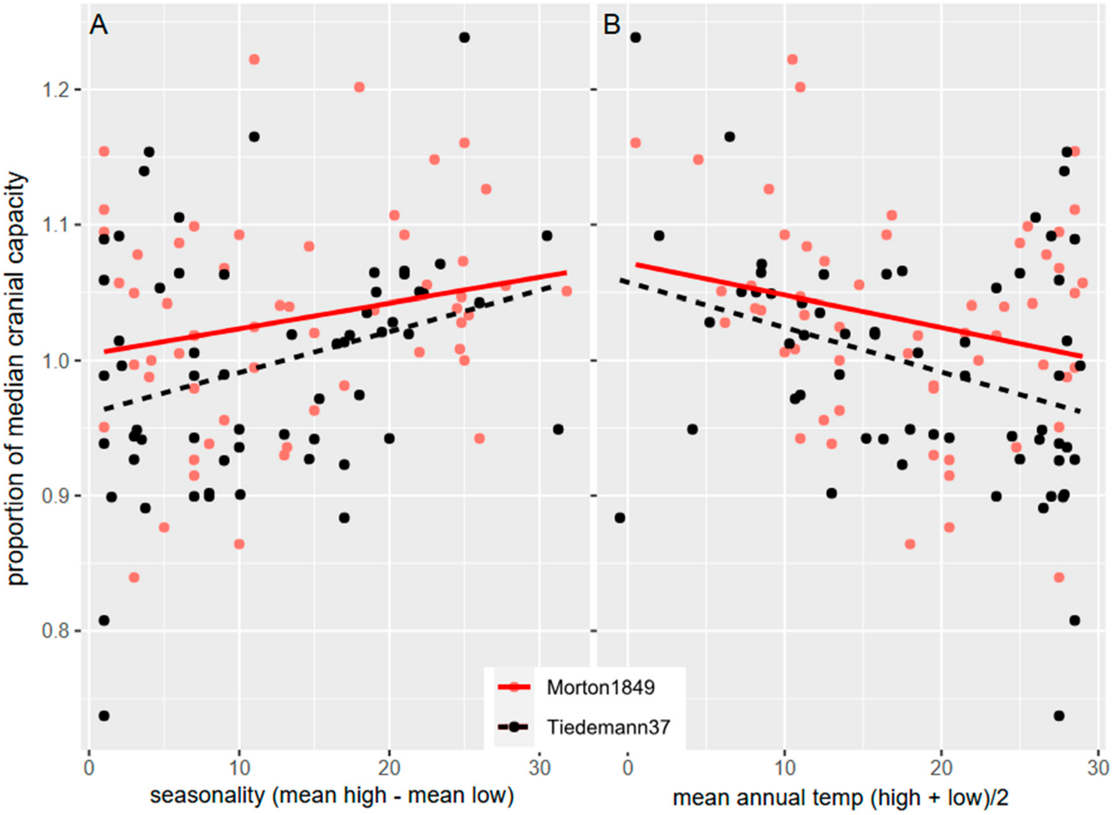

Figure 2. Cranial volumes expressed as a proportion of the median cranial capacity (y-axes) for Morton (red/light shading) and Tiedemann (black) datasets. Each point represents a mean of skulls at a given latitude, rounded to the nearest integer. X-axis shows the average of (**A**) seasonality (annual high - low temperatures) and (**B**) mean temperature ((high+low)/2) for skulls at a given degree latitude.

Using a contingency table with categories defined by median cranial size in the top or bottom 50% and latitude above or below ±45°, both Tiedemann (1837) and Morton (1849) show larger medians among the localities positioned at higher latitudes (Figure 3). Of the 17 localities in the Morton dataset at latitudes above 45°, 14 have median cranial volumes in the top half of the size range and three in the bottom half. Of the 77 localities at or below 45° latitude (north or south), 41 have medians in the bottom half and 36 in the top. For Tiedemann, 37 localities above 45° are in the top half of his size range and 17 are in the bottom. Between 45° north and 45° south of the equator, 23 median cranial volumes are in the top half and 42 are in the bottom. These patterns depart significantly from the null hypothesis of no relationship between latitude and cranial volume using either chi-squared or Fisher's exact tests (Table 3).

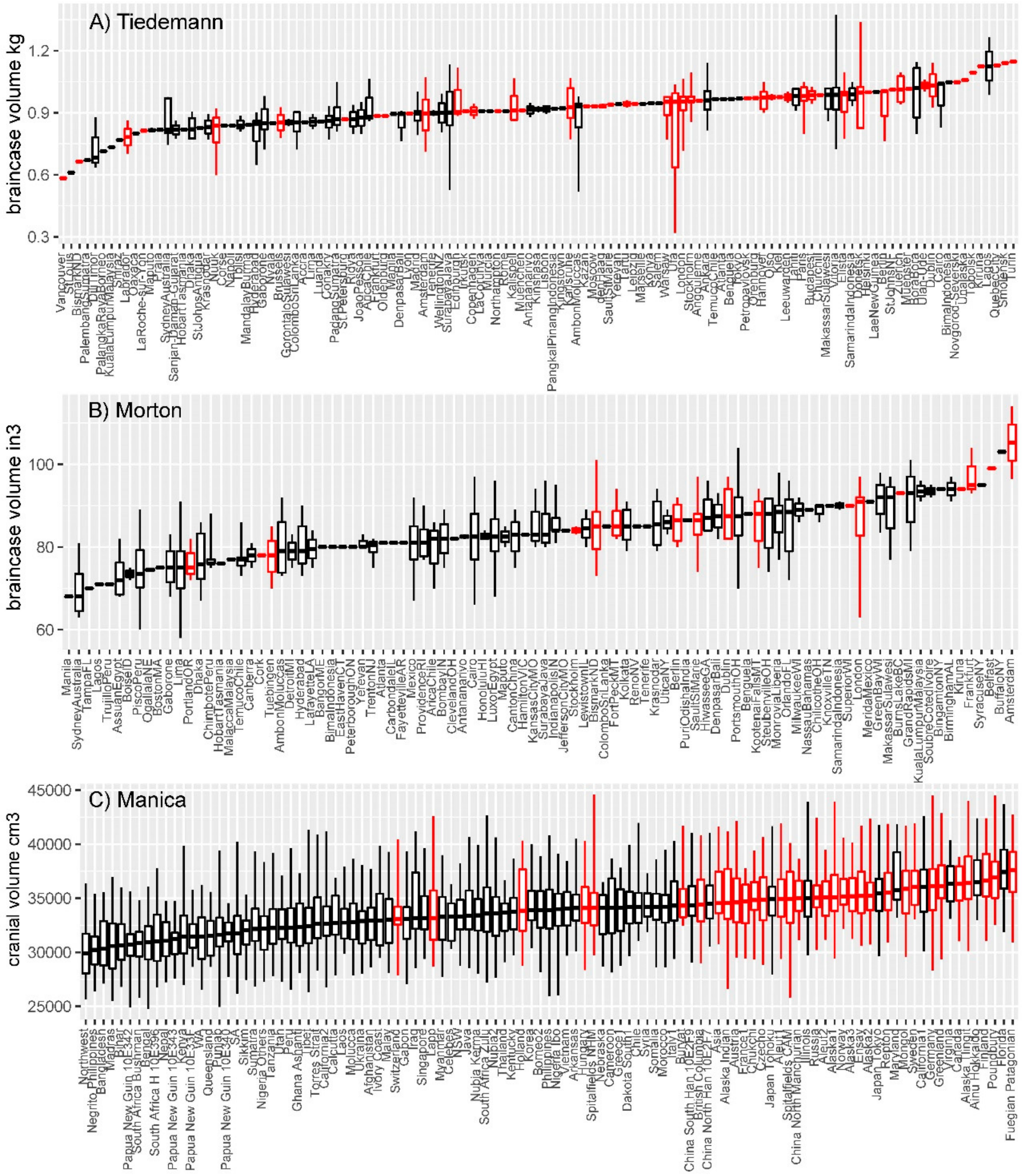

Figure 3. Indices of volume (*y*-axis) from Tiedemann (**A**, braincase in kg millet seed), Morton (**B**, braincase in cubic inches) and Manica (**C**, cranium in cubic cm, males only) by population group (*x*-axis), ordered by medians. Boxes show middle quartiles, whiskers range. Red/light-shading indicates populations from 45° or higher north or south latitudes.

Table 3. Contingency tables showing frequencies within the upper and lower 50% of cranial sizes (derived using medians of cranial size from each locality), from localities at either high (>45°) or low (<=45°) absolute degrees latitude (see Figure 3), and results of X^2 and Fisher exact tests. X^2 *p* values are derived using a Monte-Carlo simulation with 2000 replicates.

Tiedemann (1837)

	latitude >45	latitude <=45
>= 50% cranial size	37	23
< 50% cranial size	17	42

X-squared = 12.954, *p* value = **0.001**
Fisher exact test *p* value = **0.0004389**

Morton (1849)

	latitude >45	latitude <=45
>= 50% cranial size	14	36
< 50% cranial size	3	41

X-squared = 7.0882, *p* value = **0.01099**
Fisher exact test *p* value = **0.01391**

Manica et al. (2007)

	latitude >45	latitude <=45
>= 50% cranial size	30	22
< 50% cranial size	3	50

X-squared = 32.973, *p* value = **0.0005**
Fisher exact test *p* value = $\mathbf{3.559 \times 10^{-9}}$

Patterns in the Manica dataset

Data from Manica et al. (2007) also suggest that humans associated with higher latitudes tend to have larger cranial volumes (Figures 3C and 4), represented by maximum length x breadth x height (as described in Appendix 2). Again, the signal from the southern hemisphere is weaker than in the north, significant according to a Mantel test (Table 2) but not Kendall's rank correlation (Table 1). Focusing on males (as the Manica dataset lacks females at the highest southern latitudes), data from Manica are more exclusively partitioned than those from Tiedemann and Morton. There are 30 medians in the top half of the size range over ±45° and three between +45° and -45°. There are 50 medians in the bottom half between +45° and -45° and 22 above ±45°. As with the Tiedemann and Morton datasets, this distribution departs significantly from the null hypothesis of no relationship between latitude and cranial volume (Table 3).

Accordingly, cranial volumes grouped by latitude show increasing trend lines with distance from the equator (Figure 4), with the sample consisting of 50 means (one per degree latitude) representing 3609 skulls from the northern hemisphere and 18 means representing 1057 skulls from the southern hemisphere and equator. Using Kendall's rank-based test, the correlation of mean cranial volume among males with latitude is significant in the northern but not southern hemisphere (Table 1); using a Mantel test, the correlation is significant in both hemispheres (Table 2).

As noted in Appendix 2 (with raw data provided in Appendix 3), the Manica et al. (2007) dataset, derived in turn from Hanihara and Ishida (2001, 2005), listed one or more localities for each of the populations associated with their 6245 skulls, rather than identifying one locality per skull. The localities were often singular or adjacent, but in some cases could be geographically far apart. For example, they identified 209 male and 139 female skulls from "Peru" and assigned all of them a latitude of −5°. However, the localities listed under "Peru" range from −18° to −8°, from Arica (in modern Chile) to Chicama. The most extreme example of geographically distant localities within one population was for "Chile", which consisted of 16 male skulls all assigned a single latitude of −50°, but with localities ranging from "Navarin Island" (−55°) to "Iquique" (−20°).

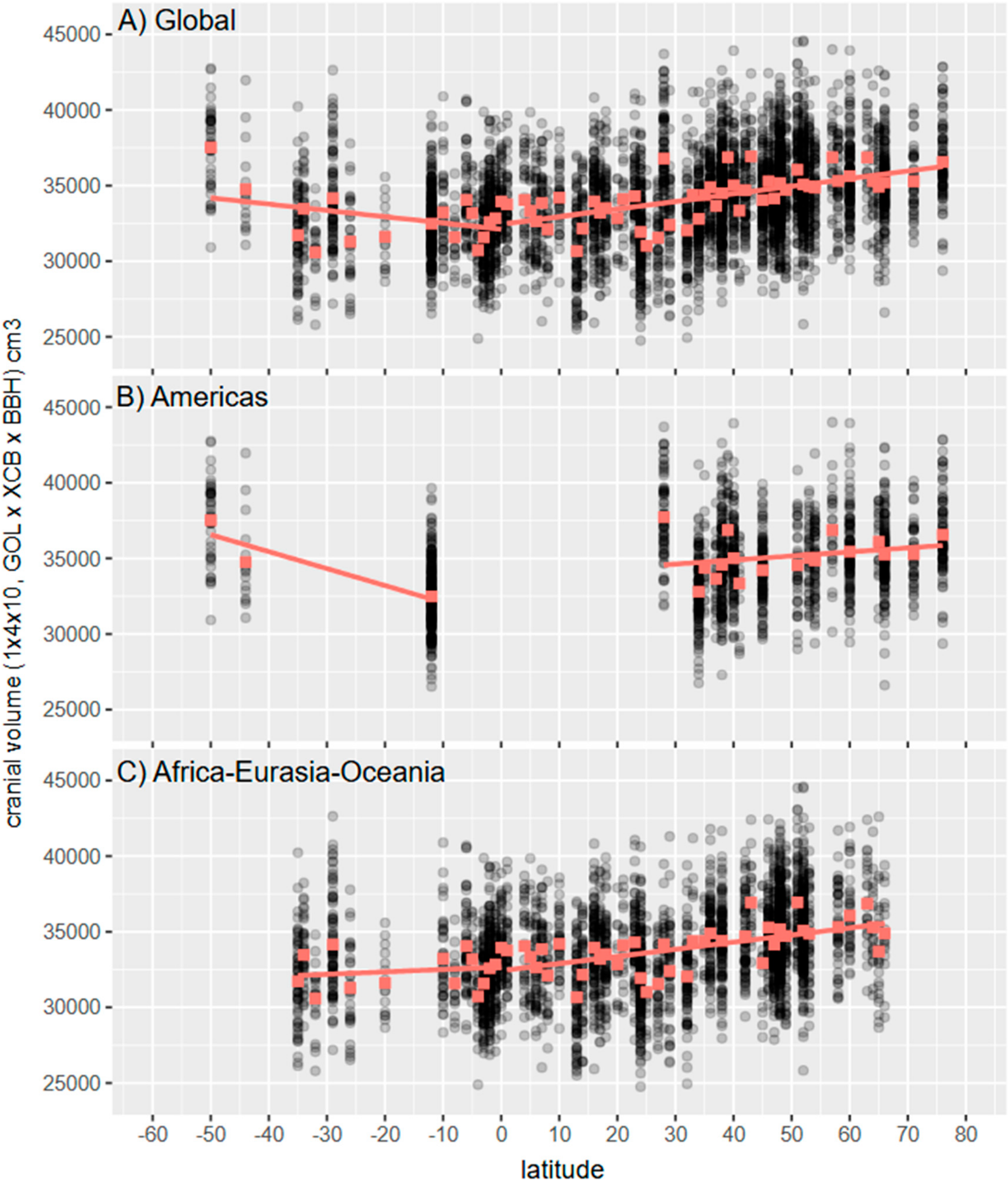

Figure 4. Cranial volume (*y*-axis) approximated by breadth x length x height in cm^3 based on measurements 1, 4, 10 of Manica et al. (2007). Circles in background represent individual male skulls; red/light-shaded squares show means for each degree latitude rounded to the nearest integer (*x*-axis).

Here, I replaced the latitudes provided by Manica et al. (2007) with the median of the localities associated with each of their populations (Appendix 3). I also undertook resampling experiments for Chile and Peru. Chile has by far the widest range (35°) of reported latitudes among its 12 localities; Peru has a smaller range (10°) but the largest number of localities (14). By randomly assigning specific latitudes corresponding to localities given by Manica et al. (2007) with individual male skulls, using 5000 replicates for each country and another 5000 for the two combined, I sought to determine how the significance values of the resulting correlation coefficients would change.

These randomizations reduced the *p* values derived from correlation tests relative to the use of medians (Figure 5). While most of the 15000 replicates remained above the significance level of 0.05 for Kendall's *Tau* (Figure 5A), nearly all (99%) dropped far below the *p* value of 0.4543 (Table 1) derived from single medians for each population. Using Pearson's *R* (which assumes a normal distribution), all of the 5000 replicates of skulls with localities in Peru resulted in *p* values under 0.05, as did 80% of the 5000 replicates for Chile and 91% of the replicates in the two combined. Nearly all of these 15000 resamples were below the non-significant *p* value of 0.103 for Pearson's *R* based on single median latitudes for the Chile and Peru populations (Figure 5B).

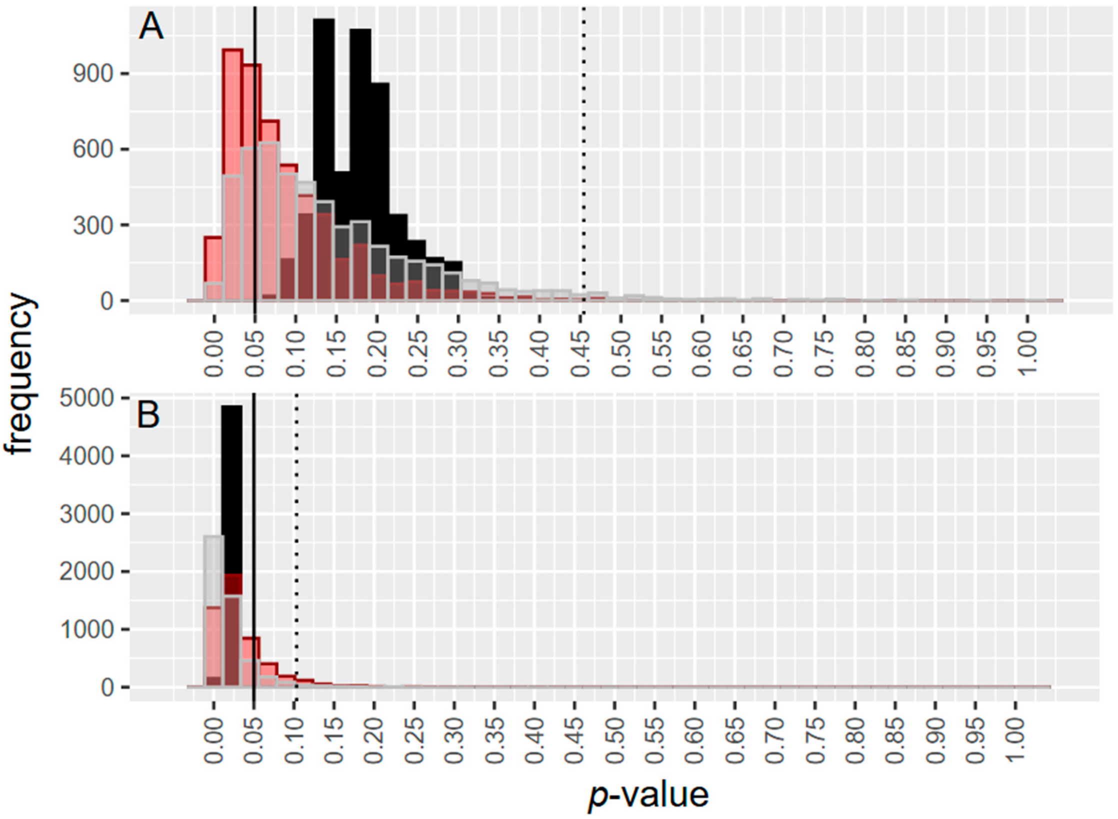

Figure 5. Histogram showing frequency (*y*-axis) of *p* values (*x*-axis) derived from replicates of Kendall's Tau (**A**) and Pearson's *R* (**B**) between latitude and cranial volume in the southern hemisphere for the Manica dataset. In each replicate, crania were randomly assigned to one of the known latitudes within Chile (red/dark shading), Peru (black), or both (light gray) with 5000 iterations each. Vertical dotted lines show the *p* value derived from a single median of latitudes within each country; solid vertical line shows significance at alpha = 0.05.

A reason for a weak signal in the southern hemisphere is the fact that the habitable, high-latitude south is tiny compared to that of the north. The only groups sampled over 43° south are Fuegians and Chileans, compared to 35 such groups over 43° in the northern hemisphere. Accordingly, without data from the Americas (Figure 4C), the southern hemisphere trendline for cranial volume is relatively flat, with no clear changes according to latitude. In addition, the Manica dataset lacks skulls from low latitudes (ca. 0–25°) in the Americas, and the Americans with the largest mean cranial volume are Seminoles (associated with the southeastern USA), with a median latitude of just 28° north. This reduces the slope of the best-fit line in the American northern hemisphere (Figure 4B) relative to that for the rest of the world (Figure 4A). The correlation between latitude and cranial volume for the American north is still positive, but lacks statistical significance.

Uneven geographic sampling and northern bias

Both datasets published in the 19th century have a much smaller sample of high-latitude, southern individuals compared to the Manica dataset, which is itself limited by the small number of habitable, high southern latitudes compared to those in the north. The most southerly locality present in Tiedemann and Morton was Tasmania (42.9° south). Only two other Tiedemann individuals (from New Zealand) and none from Morton exceeded 40° south. Just 24 of 880 skulls (3%) in the 19th-century datasets were from localities over 30° south, in contrast to 356 (40%) over 30° and 312 (34%) over 40° degrees north. Distinct results from north vs. south are consistent with Foster and Collard (2013: 13), who report that "unless a region spans more than 50 degrees of latitude or has a range of mean annual temperature in excess of 30°C, it is unlikely that a correlation between body size and latitude/mean annual temperature is reliable evidence for ... Bergmann's rule." In both the Morton and Tiedemann datasets, only North American and Asian localities show differences in annual maximum vs. minimum temperatures in excess of 30°, and only the northern hemisphere is sampled by more than 50° latitude.

The Manica dataset also has fewer individuals from the south compared to the north, and is furthermore male biased, as all of the skulls from >50° south are male. Nonetheless, for males at any locality worldwide, the "Fuegian Patagonian" median cranial volume is the largest (Figure 3C).

Discussion

Is there evidence for latitudinal change in 19th-century data on humans?

The goal of this study is to test if data on human cranial capacity from Tiedemann (1837) and Morton (1849) reflect patterns evident in more recent publications (e.g., Beals et al. 1984; Gustafsson and Lindenfors 2009; Katz et al. 2016), in which size correlates with latitude. This in turn might have led these 19th-century investigators to consider the latitude–body size mechanism articulated by Bergmann. With some caveats, data from both Morton and Tiedemann show a correlation between size and latitude.

Neither 19th-century investigator had adequate samples from the southern hemisphere to have been able to recognize latitudinal gradients there (cf. Foster and Collard 2013). However, both had sufficient samples from the north to show a significant, upward trending pattern of cranial capacity with increasing latitude, although this would not have been evident to Tiedemann had he examined his American samples alone (Figure 1B). Based on these 19th-century measurements of human skulls, it would have been difficult, but still possible, to recognize at least some aspects of increasing size with distance from the equator, as articulated by Bergmann. Data from both Tiedemann and Morton exhibited increasing cranial capacities with increasing northern latitudes (Figure 1; Tables 1 and 2), with larger crania disproportionately represented among higher latitudes (Figure 3, Table 3). Both datasets also tended to associate larger cranial capacities with localities showing lower average temperatures and more disparate extremes of seasonal temperatures (Figure 2), although non-parametric, rank-based tests showed significant relationships only for the Tiedemann dataset (Table 1).

Historical background

It is worth considering the context in which Tiedemann undertook his investigation, and point out ways in which he differed from contemporaries such as Samuel Morton. Tiedemann (1837) started his monograph noting the relevance of comparative anatomy to debates in the British parliament concerning the slave trade, and made no secret of his antipathy to slavery. He included numerous tables of raw data and summaries thereof, all of which are meant to answer two questions he set out in his introduction: "I) Is the brain of [sub-Saharan Africans] different from that of Europeans? II) Is the brain of [sub-Saharan Africans] more similar to that of *[Pongo pygmaeus]* than Europeans?" Phrasing research questions in this way is of course offensive today, but it would be unfair to impose our standards on a naturalist writing over 180 years ago. Tiedemann analyzed not only hundreds of skulls, but also 65 human cadavers of European origin and a number of juvenile and infant skulls, with which he made comparisons to the few sub-Saharan African cadavers he had. Regarding actual brain weights (augmenting his approximations thereof based on cranial capacity; see Tobias 1970), Tiedemann wrote (1837: 20) "the claim of many naturalists and anatomists that [sub-Saharan Africans] have smaller brains than Europeans can under no circumstances be regarded as proven". Regarding comparisons of soft-tissue morphology, Tiedemann wrote (1837: 55) "... the spinal cord of the [sub-Saharan African] Honore ... shows no major differences from that of the European. ... [His] cerebellum is in its outer form, sulci, divisions and lobes completely equivalent to that of the European". Tiedemann (1837: 61) also disputed a claim that sub-Saharan Africans had thicker cranial nerves.

Tiedemann did not claim that all aspects of human brain morphology were identical across the groups he examined. Concerning the four brains of individuals from sub-Saharan Africa at his disposal, he wrote that the "sulci and gyri of both hemispheres of the cerebrum ... exhibited a more symmetrical arrangement than what one tends to perceive from the brain of Europeans" (1837: 59), but expressed doubt that this was a consistent feature (1837: 63). As pointed out by Richards (2018), Tiedemann also acknowledged a more frequent occurrence of unusually large cranial capacities among his "Caucasian" and "Malay" samples. Foreshadowing Gould's discussion of population differences some 150 years later (see below), Tiedemann was aware of allometry and suggested differences in overall stature were behind differences in cranial data. Overall, Tiedemann accurately expressed an anatomical basis for his conviction that human brains were essentially similar in their anatomy and not derived from distinct species.

However, he referred to latitude in his text only once (1837: 50) to orient his discussion of variation among African populations. Nowhere does he mention the pattern, subtle but evident in his own data, that skulls from high Eurasian latitudes tended to have larger cranial capacities than those from low latitudes (Figure 1, Figure 3A), and relatedly that those from cooler, more seasonal parts of the world tended to be larger (Figure 2; Table 1). He also did not recognize that his own measurements of 141 European cranial volumes were, on average, significantly larger than the 35 he had from sub-Saharan Africa (Table 4A).

Tiedemann wrote a decade before Bergmann (1848), considered soft tissues as well as skulls, took allometry into consideration, and discovered that human brains are very similar despite the different geographic origins of their owners. For his time, Tiedemann was a visionary. His approach underscores a willingness to explore socially divisive questions with data he thought relevant from the natural world, and he did not shy away from conclusions that would have been unpopular for his time (Richards 2018). He can hardly be faulted for not getting everything right about craniometric variation. Nonetheless, his focus on the sociocultural implications of “race” appears to have precluded discovery of geographic differences that were at least partly evident in his own data.

Samuel Morton’s 1849 book consisted almost entirely of a list of 1512 human skulls (514 of which had sufficient metadata to be sampled here), with “F.A.” (facial angle) and “I.C.” (internal capacity, measured in cubic inches) for most, along with details on their identity and origins. The only substantial text consists of about five pages in his introduction. Morton ignored soft-tissue anatomy, focused only on cranial capacity with its implications for brain size, assumed that larger equaled more intelligent and did not consider relative size (in contrast to Tiedemann). Morton’s mistaken generalizations were published in his earlier work, such as his view that human ethnic groups were of independent origins (see Richards 2018: 154). Like Tiedemann, Morton also attempted to apply insights from the natural world to social questions of his time; he too missed the correlation of increasing size with latitude in the northern hemisphere (Figures 1 and 3; Tables 1–3).

Table 4. Significance values for pairwise Wilcoxon Rank Sum tests of differences between proportion median skull sizes among geographic units for (A) Tiedemann, (B) Morton, and (C) Manica datasets, using a Bonferroni correction. Underlined values indicate significant differences within dataset at alpha < 0.05. Bold values indicate agreement of a 19th-century dataset with that of Manica (which lacks samples from Pacific and central American regions).

(A) Tiedemann proportion median skull size

	Asia	AustralNG	CAmerica	Easia	Europe	NAfrica	NAmerica	Pacific	SAmerica	Sasia	SEasia
AustralNG	1.0000	-	-	-	-	-	-	-	-	-	-
CAmerica	1.0000	1.0000	-	-	-	-	-	-	-	-	-
Easia	1.0000	1.0000	1.0000	-	-	-	-	-	-	-	-
Europe	1.0000	1.0000	1.0000	1.0000	-	-	-	-	-	-	-
NAfrica	1.0000	1.0000	1.0000	1.0000	1.0000	-	-	-	-	-	-
NAmerica	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-	-	-	-	-
Pacific	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-	-	-	-
SAmerica	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-	-	-
Sasia	**0.0056**	1.0000	1.0000	1.0000	**1.1e-05**	1.0000	1.0000	0.0284	0.4133	-	-
SEasia	1.0000	1.0000	1.0000	1.0000	0.7617	1.0000	1.0000	1.0000	1.0000	0.4657	-
subsAfri	0.6115	1.0000	1.0000	1.0000	**0.0110**	1.0000	1.0000	0.6332	1.0000	1.0000	1.0

(B) Morton proportion median skull size

	Asia	AustralNG	CAmerica	Easia	Europe	NAfrica	NAmerica	Pacific	SAmerica	Sasia	SEasia
AustralNG	1.00000	-	-	-	-	-	-	-	-	-	-
CAmerica	1.00000	1.00000	-	-	-	-	-	-	-	-	-
Easia	1.00000	1.00000	1.00000	-	-	-	-	-	-	-	-
Europe	1.00000	**0.04152**	0.10712	1.00000	-	-	-	-	-	-	-
NAfrica	1.00000	1.00000	1.00000	1.00000	0.02286	-	-	-	-	-	-
NAmerica	1.00000	0.09909	1.00000	1.00000	1.00000	0.49386	-	-	-	-	-
Pacific	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	-	-	-	-
SAmerica	0.11203	1.00000	0.01221	1.00000	**4.7e-11**	**8.8e-09**	**< 2e-16**	0.91330	-	-	-
Sasia	1.00000	1.00000	1.00000	1.00000	**0.00940**	1.00000	0.13270	1.00000	0.04336	-	-
SEasia	1.00000	0.56485	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	**1.0e-04**	1.0	-
subsAfri	1.00000	0.40765	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.00013	1.0	1.0

(C) Manica proportion median skull size

	Asia	AustraliaNG	Easia	Europe	NAfrica	NAmerica	SAmerica	Sasia	SEasia
AustralNG	< 2e-16	-	-	-	-	-	-	-	-
Easia	1.00000	< 2e-16	-	-	-	-	-	-	-
Europe	1.00000	< 2e-16	1.00000	-	-	-	-	-	-
NAfrica	1.00000	7.7e-16	1.00000	1.00000	-	-	-	-	-
NAmerica	0.00091	< 2e-16	2.7e-06	1.2e-10	1.00000	-	-	-	-
SAmerica	5.1e-15	1.00000	< 2e-16	< 2e-16	1.8e-08	1.1e-11	-	-	-
Sasia	< 2e-16	1.00000	< 2e-16	< 2e-16	< 2e-16	< 2e-16	0.32707	-	-
SEasia	0.00211	< 2e-16	1.3e-05	1.3e-06	1.00000	1.00000	1.0e-08	< 2e-16	-
subsAfri	9.6e-12	2.3e-06	< 2e-16	< 2e-16	2.2e-05	3.1e-06	0.17267	7.8e-08	0.00039

Morton’s actual measurements of cranial volumes contradicted the conclusions about cranial size across human groups often attributed to him. His samples were limited, with just 21 crania from sub-Saharan Africa and 37 from Europe (Appendix 3), but his own measurements showed no significant difference between the two groups (Table 4B). The much larger sample of Manica et al. (2007) included 631 crania from sub-Saharan Africa and 1092 from Europe. These and most other continentally-defined human populations differ significantly in skull size, and European crania are significantly larger than those from other regions except for Asia, east Asia, and north Africa (Table 4C).

Cranial size and "intelligence"

Although cranial size and shape exhibit developmental plasticity, many aspects of cranial anatomy are heritable (Relethford 2004), and brain size shows evolutionary variation throughout Mammalia (Mace et al. 1981; Burger et al. 2019). This variation has implications for the evolution of at least some aspects of "intelligence" (Van Valen 1974), and there are indeed correlations of physical traits with performance on intelligence tests, including brain volume (Ritchie et al. 2015) and myopia (Williams et al. 2017). There are also differences between human sexes in the density of neural connections and cortical size variation, correlated at least partly with fetal exposure to sex hormones (Gilmore et al. 2018; Baron-Cohen 2020). However, osteological cranial capacity is not a straightforward representation of brain mass (Tobias 1970) and some of history's most intellectually gifted humans had below-average brain weights (Witelson et al. 1999). Furthermore, the close correlation of stature with brain size (Tobias 1970; Beals et al. 1984; Skullerud 1985) means that, in absolute terms, heavyweight professional wrestlers will on average have larger brains than mathematicians, as will Pleistocene compared to modern humans (Henneberg 1988), or males compared to females. These observations suggest that raw cranial capacity, within the overall range known for humans, is not a consistent indicator of "intelligence".

Stephen Jay Gould and the 20th century

The preoccupation with cranial variation among human populations and its supposed relation to intelligence and sociocultural value has persisted in the scientific literature well beyond the 19th century, although a few investigators did investigate climatic factors (e.g., Thomson 1913; Roberts 1954; Coon 1955). Probably the most influential critique of human craniometrics was Gould's *The Mismeasure of Man*, first published in 1981, reprinted in 1996, and preceded by Gould's 1978 article in *Science* on Samuel Morton. Gould (1978, 1996) argued that Morton (1839, 1849) was biased in arguing for larger brain size among Europeans over other human groups, according to the dominant prejudices of his time (Richards 2018). As Gould observed, Morton's work was used throughout the 19th century to justify white supremacy. Gould scrutinized how Morton (1839, 1849) presented his data, and found his work to be a "patchwork of fudging and finagling" (1996: loc. 1323). However, Gould's work has in turn been evaluated to reveal his own bias (Michael 1988; Lewis et al. 2011; Richards 2018; see also Mitchell 2018), which Gould partly acknowledged in a footnote to his 1996 reprint (1996: loc. 8218).

Gould lauded Tiedemann's argument for racial equality, which for both meant demonstrating similarity among human ethnic groups, and noted (Gould 1996: loc. 1752) an old critique of Tiedemann: the imprecision of millet seeds to measure cranial capacity. Gould reasoned that given more- (seed) and less- (metal shot) error-prone means of measuring cranial capacity, differences between the two showed that Morton underestimated the capacities of African crania more than those of other groups (see Weisberg and Paul 2016).

However, there are several mistakes in Gould's conclusions. Firstly, an assistant of Morton made the relevant seed measurements, not Morton, which were in any event scrapped when Morton became suspicious of the errors (Richards 2018: 173). Secondly, Mitchell (2018) noted that Morton's skulls measured for the 1849 book with shot were probably not the same those measured for the 1839 book with seed, and that "errors in the seed measurements were significant, but likely random" (Mitchell 2018: 7). In other words, they were not distributed in such a way as to portray sub-Saharan Africans with unduly small cranial capacities.

Morton did project his assumptions about ethnicity on at least some of his data. For example, and as noted by Gould (1996: loc. 8412), he apparently used large size as a criterion to group certain Egyptian skulls as "Pelasgic" (i.e., of Greek ancestry). Such specimens do indeed have a significantly larger average cranial capacity than others from Egypt. In contrast, Egyptian skulls labeled by Morton (1849) as "negroid" (which he presumed to be of mixed ancestry) are not significantly different than non-"Pelasgic" skulls from the same localities. Moreover, actual re-measurements of the skulls in Morton's collection by two different sets of investigators (Michael 1988; Lewis et al. 2011) are not consistent with Gould's accusation of pervasive, unconscious bias. Lewis et al. (2011: 5) concluded that the "summary table of Morton's final 1849 catalog ... has multiple errors .... However, had Morton not made those errors his results would have more closely matched his presumed a priori bias .... Ironically, Gould's own analysis of Morton is likely the stronger example of a bias influencing results."

Lewis et al. (2011) were obviously not arguing for any of the repugnant racial stereotypes favored by Morton. Instead, they took issue with Gould's claim that Morton unconsciously manipulated his data to favor white supremacy. They did not discuss what the population-level differences in Morton's data might mean. However, unlike Gould, Lewis et al. (2011) did cite past work on latitude and temperature (Beals et al. 1984) as a potential explanation. Gould (1996) gave examples from various 19th-century writings about changes in human skin color and character given extended periods of time in a non-native climate, and even had an index entry entitled "climate, racial differences attributed to". However, he did not consider latitudinal variation in size, which by the 1970s was a reasonably well-established pattern (Rensch 1924, 1938; Mayr 1942, 1956; Lindsey 1966; Beals 1972), albeit with caveats about cause and effect (e.g., James 1970). Gould himself discussed examples of Bergmann's rule in his 1977 book *Ever Since Darwin* and in a 1972 review paper on geographic variation. Gould and Johnston (1972) noted instances in the literature for latitudinal patterns consistent with Bergmann's rule in birds, deer, and ticks, but not humans. They concluded that "we know that Bergmann's rule governs the distribution of body size in many homeothermic species" (Gould and Johnston 1972: 488) and underscored the importance of quantification.

Yet Gould did not mention Bergmann's rule as a potential factor behind human size variation in his 1978 paper, or in the 1981 or 1996 editions of *The Mismeasure of Man*. According to Gould, "Morton assumed that [cranial capacity indicates] an innate difference in intelligence. He never considered any other proposal for the disparity in average cranial capacity – though another simple and obvious explanation lay before him" (1996: loc.1403) which, according to Gould, was stature: "size of brains are related to the sizes of bodies that carry them". In other words, differences in brain size are due to size. In an unusual lapse of curiosity, Gould (1996) did not proceed to ask why the "sizes of bodies" would themselves have varied across human populations, despite having noted the mechanism articulated by Bergmann (1848) in his own previous writings (e.g., Gould and Johnston 1972; Gould 1977).

Some later investigators (Weisberg 2014; Weisberg and Paul 2016; Mitchell and Michael 2019) have sought to validate Gould's noble, egalitarian sentiment as expressed in *The Mismeasure of Man* by arguing that Gould's central claim of Morton's bias is still valid. All of these authors are correct to express disdain for Morton's racism, which is amply demonstrated by his own publications; its recognition this does not require any reevaluation of Morton's methods for collecting or analyzing data (Richards 2018). Perhaps more interestingly, it was not Morton but Gould who selectively presented his data to fit his preconceived bias (Lewis et al. 2011), i.e., that there should not be variation among human ethnic groups. Among these recent publications favorable to Gould, Weisberg (2014: 177) was among the few to mention Bergmann (1848) as relevant to human size variation, but still agreed with Gould that "there are no significant differences in mean cranial capacities across races in Morton's collection".

Categorizing human variation

Templeton (1998, 2013) argued that, if "race" is supposed to mean a biological subspecies, then there are two main reasons why human "races" do not exist: 1) lack of genetic differentiation (i.e., F_{st} values below 0.25) and 2) the absence of a bifurcating, tree-like structure in the Pleistocene history of *Homo sapiens*. Compared to other primates, humans have relatively little genetic diversity (Osada 2015), and most diversity that does exist is distributed within human groups rather than between them (Barbujani et al. 1997). However, measuring human genetic diversity using F_{st} is complex and not well-suited to human geographic variability (Long and Kittles 2003; Roseman 2021). Edwards (2003) also criticized the use of F_{st} in assessing genetic diversity within and between human ethnic groups. However, unlike Long and Kittles (2003) and Roseman (2021), Edwards (2003) viewed the existence of those groups as at least partly coincident with their genetic differentiation. Perhaps more importantly, and regardless of how one interprets F_{st} (see Novembre 2022), Edwards (2003: 801) warned that "it is a dangerous mistake to premise the moral equality of human beings on biological similarity because dissimilarity, once revealed, then becomes an argument for moral inequality." A similar point was made by Reich (2018: 257–258): "we should prepare our science and our society to be able to deal with the reality of differences [between human populations] instead of sticking our heads in the sand and pretending that differences cannot be discovered." Both Edwards (2003) and Reich (2018) address an issue which is obfuscated by Gould and others who ignore or downplay variation between human groups: our "moral equality" (Edwards 2003: 801), i.e., our status as deserving of equality of opportunity and social justice, is contingent upon our shared dignity as human beings, not upon claims of phenotypic or genotypic sameness.

Less controversial is the understanding that the amount of genetic exchange among humans worldwide has been substantial for millennia (Templeton 2013; Reich 2018). Depictions of tree-like evolution among modern humans mask some of the recursive or "trellis"-like patterns that actually occurred (Templeton 2013: fig. 2). On this basis alone, humans cannot be legitimately categorized into taxonomic subspecies. Just as there is no evidence to partition human groups into biological subspecies, there is no evidence for claims of superiority of one ethnic group over another. This observation does not mean that human groups have to be phenotypically or genetically "equal". In fact, a nuanced understanding of human diversity recognizes that "hereditary differences between human individuals are real and important, and there are significant average differences in various respects between some regional populations" (Cartmill 1998: 659). Such differences may result from non-random mating (Burrell and Disotell 2009) and subsequent departures from Hardy-Weinberg equilibrium for certain genetic loci (Sebro et al. 2017). There are also minor differences in morphology between populations (Howells 1995; Ousley et al.

2009; Yong et al. 2018). Recognition of such population-level genetic or morphological differences does not imply approval of the term "race", which carries with it many social and cultural factors that have little connection to biological heredity (Templeton 2013). Indeed, many authors (e.g., Cartmill 1998) have argued that the negative, sociocultural baggage of that term warrants its exclusion in discussions of human biological diversity.

According to Wagner et al. (2017), a majority (86%) of professional anthropologists disagreed with their survey question that "the human population may be subdivided into biological races" (Wagner et al. 2017: table 2, question #1), but there is less concordance on the relevance of geography to human variation, which does have a demonstrable association with indices of genetic and phenotypic distance (Manica et al. 2005, 2007). The anthropological community seems more divided regarding the survey question of Wagner et al. (2017: table 2, question #7) that "continental population categories – Africans, Asians, Europeans – are useful for examining genetic relationships (i.e., relatedness) among people". 38% of surveyed anthropologists disagreed, but 33% agreed. Given its political baggage, "race" might therefore reasonably be replaced with alternatives such as "ethnic group" (Montagu 1964) or "phyle" (Sánchez-Villagra 2022).

All of these terms represent non-exclusive, relatively open units of gametic exchange across humanity. Whichever one is used, the challenge is to convey simply and uncontroversially the ways in which some human ethnic groups show, on average, significant differences from one another, including cranial size and shape. For example, Fuegian Patagonians have significantly larger cranial volumes than Papuans (Figure 3C). These groups exhibit substantially different histories and are clearly distinct ethnic groups, even though neither one (or any other two modern human populations) exhibits sufficient differentiation to qualify as distinct biological subspecies.

Gould (1978: 503) claimed that "When [Morton's] data are properly reinterpreted, all [human] races have approximately equal capacities". However, he did so not on the basis of craniometric data, but presumably due to his desire to support human equality in spite of persistent socioeconomic and political inequality. In fact, cranial capacities between several geographically distinct human groups – some of which are similar to the "races" of Tiedemann and Morton – are not "approximately equal" (Table 4). They are in terms of intellectual potential, but there are at least some phenotypic differences (Figure 3C, Table 4; see also, e.g., Howells 1995; Ousley et al. 2009; Yong et al. 2018). Ironically, Gould's assertion about the variation of cranial capacity among humans, when shown to be false, may lead to skepticism about his other assertions, including the rightful pursuit of socioeconomic justice. Human populations should not have to be "approximately equal" in a phenotypic or genotypic sense to qualify for social justice or equality of opportunity.

Conclusions

Nineteenth-century naturalists did not know that some indigenous groups of the African continent (e.g., Khoisan) harbor by themselves more genetic variation than all other human groups combined (Schlebusch et al. 2020), that they also exceed non-African groups in their phenotypic diversity (Manica et al. 2007), or that modern social constructs of ethnicity do not correspond neatly with biological categories based on ancestry. Nor did they know that all non-African humans derive from a limited number of colonization events from Africa, with some hybridization with archaic groups such as Denisovans and Neanderthals, over the last 100,000 years (Stringer 2014; Sankararaman et al. 2016). Preoccupation with the implications of craniometrics for human variation, particularly as it concerns intelligence, was another factor that may have prevented investigators such as

Morton, Tiedemann, and more recently Gould, from fully appreciating differences among human populations. Of course, Gould and Tiedemann stand in stark contrast to Morton in terms of their accurate recognition of the unity of the human species. Nonetheless, these authors seemed not to have noticed among humans a pattern that Bergmann (1848) recognized among other vertebrates: a correlation between size and latitude.

Investigators without access to evolutionary or statistical methods (particularly Tiedemann, who got so much else right) can be forgiven. A more recent investigator such as Gould, who had even published on Bergmann's pattern (Gould and Johnston 1972), could have acknowledged it in *The Mismeasure of Man* and in his 1978 article in *Science*, but did not. By the end of the 20th century, and notwithstanding the genuine uncertainty that persists regarding the importance of climate relative to other causative factors, there was a reasonably clear, empirical basis for latitudinal size gradients in multiple vertebrate groups (Lindsey 1966), including humans (e.g., Beals et al. 1984). At least in print, Gould (1978, 1996) seemed not to have noticed such gradients in humans, possibly due to his preoccupation with the sociocultural implications of human "race".

Acknowledgments: I am grateful to many colleagues and reviewers who provided helpful critiques of past versions of this manuscript. I have chosen not to name them here due to the controversial nature of the article and my desire not to implicate valued colleagues. A brief history of the review process for this manuscript might help to convey the nature of this controversy. A previous version of this paper was submitted to "Journal A" in June of 2021. It received two constructive reviews along with the subject-editor's recommendation of "minor revision" in late July 2021. One or more members of the journal's "managing board" then intervened and asked the subject-editor to send it out for additional review, which I received in October of 2021. This review was generally positive: "Overall, I think the paper makes a good contribution to the literature in its findings of the pattern of cranial size, and its demonstration that old data are still usable. However, I do have some suggestions that could strengthen the paper." I accordingly sent another revision to Journal A, which by this time was being assessed by the journal's "managing board" who rejected the revised manuscript in late November 2021. I then submitted the manuscript to "Journal B" in early February of 2022. Thus began a process lasting 1.7 years until late October 2023 when the paper was finally rejected, spanning two senior editors, about six reviewers, and five revisions after the initial submission, two of which resulted in a decision of "minor revision". After the third submission in January 2023, the then-senior editor wrote, "The reviewers who were white men either suggested minor revisions or accepted the revisions you implemented. In contrast, the reviewers who were women and/or people of color were far more critical. This places me in a challenging position ... ." At this point I should have realized that current politics would outweigh my further attempts to craft the text to the satisfaction of all referees at this publication. I am grateful to *The Journal of Controversial Ideas* for the opportunity to finally publish this article.

Supplementary Materials: The following supporting information can be downloaded here, **Appendix 1** is the original German from Tiedemann (1837) and Bergmann (1848) cited in the text and translated into English by the author. **Appendix 2** is an expanded description of methods. **Appendix 3** is a zip file with raw data on cranial volume and geography derived from Tiedemann (1837), Morton (1849) and Manica et al. (2007) in one spreadsheet ("comb.csv"). "ManicaAll.ods" contains the Manica dataset with the localities used to calculate median latitudes per population.

References

Ashton, K.G., Tracy, M.C., and Queiroz, A. de. 2000. Is Bergmann's rule valid for mammals? *American Naturalist* 156: 390–415.

Barbujani, G., Magagni, A., Minch, E., and Cavalli-Sforza, L.L. 1997. An apportionment of human DNA diversity. *Proceedings of the National Academy of Sciences* 94: 4516–19.

Baron-Cohen, S. 2020. *The Pattern Seekers: How Autism Drives Human Invention*. Dublin: Basic Books.

Beals, K.L. 1972. Head form and climatic stress. *American Journal of Physical Anthropology* 37: 85–92.

Beals, K.L., Smith, C.L., and Dodd, S.M. 1983. Climate and the evolution of brachycephalization. *American Journal of Physical Anthropology* 62: 425–37.

Beals, K.L., Smith, C.L., Dodd, S.M., Angel, J.L., Armstrong, E., Blumenberg, B., Girgis, F.G., Turkel, S., Gibson, K.R., Henneberg, M., et al., 1984. Brain size, cranial morphology, climate, and time machines [and comments and reply]. *Current Anthropology* 25: 301–30.

Bergmann, C. 1848. *Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse*. Göttingen: VandenHoeck & Ruprecht.

Bogin, B., Hermanussen, M., and Scheffler, C. 2022. Bergmann's rule is a "just-so" story of human body size. *Journal of Physiological Anthropology* 41: 15.

Burger, J.R., George, M.A., Jr., Leadbetter, C., and Shaikh, F. 2019. The allometry of brain size in mammals. *Journal of Mammalogy* 100: 276–83.

Burrell, A.S., and Disotell, T.R. 2009. Panmixia postponed: ancestry-related assortative mating in contemporary human populations. *Genome Biology* 10: 1–4.

Cartmill, M. 1998. The status of the race concept in physical anthropology. *American Anthropologist* 100: 651–60.

Coon, C.S. 1955. Some problems of human variability and natural selection in climate and culture. *American Naturalist* 89: 257–79.

Edwards, A.W.F. 2003. Human genetic diversity: Lewontin's fallacy. *BioEssays* 25: 798–801.

Foster, F., and Collard, M. 2013. A reassessment of Bergmann's rule in modern humans. *PLoS ONE* 8: e72269.

Freckleton, R.P., Harvey, P.H., and Pagel, M. 2003. Bergmann's rule and body size in mammals. *American Naturalist* 161: 821–25.

Geist, V. 1987. Bergmann's rule is invalid. *Canadian Journal of Zoology* 65: 1035–38.

Gigliotti, L.C., Berg, N.D., Boonstra, R., Cleveland, S.M., Diefenbach, D.R., Gese, E.M., Ivan, J.S., Kielland, K., Krebs, C.J., Kumar, A.V., et al., 2020. Latitudinal variation in snowshoe hare (Lepus americanus) body mass: a test of Bergmann's rule. *Canadian Journal of Zoology* 98: 88–95.

Gilmore, J.H., Knickmeyer, R.C., and Gao, W. 2018. Imaging structural and functional brain development in early childhood. *Nature Reviews Neuroscience* 19: 123–37.

Gould, S.J. 1977. *Ever since Darwin: Reflections in Natural History*. New York: WW Norton.

Gould, S.J. 1978. Morton's ranking of races by cranial capacity. Unconscious manipulation of data may be a scientific norm. *Science* 200: 503–509.

Gould, S.J. 1996 (first printed 1981). *The Mismeasure of Man*. New York: WW Norton.

Gould, S.J., and Johnston, R.F. 1972. Geographic variation. *Annual Review of Ecology, Evolution, and Systematics* 3: 457–98.

Gustafsson, A., and Lindenfors, P. 2009. Latitudinal patterns in human stature and sexual stature dimorphism. *Annals of Human Biology* 36: 74–87.

Hanihara, T., and Ishida, H. 2001. Os incae: variation in frequency in major human population groups. *Journal of Anatomy* 198: 137–52.

Hanihara, T., and Ishida, H. 2005. Metric dental variation of major human populations. *American Journal of Physical Anthropology* 128: 287–98.

He, J., Tu, J., Yu, J., and Jiang, H. 2023. A global assessment of Bergmann's rule in mammals and birds. *Global Change Biology* 29: 5199–210.

Henneberg, M. 1988. Decrease of human skull size in the Holocene. *Human Biology* 60: 395–405.

Howells, W.W. 1995. Who's who in skulls: ethnic identification of crania from measurements. In *Papers of the Peabody Museum of Archaeology and Ethnology*. Cambridge MA: Harvard University Press, Vol. 82.

Hubbe, M., Hanihara, T., and Harvati, K. 2009. Climate signatures in the morphological differentiation of worldwide modern human populations. *Anatomical Record* 292: 1720–33.

James, F.C. 1970. Geographic size variation in birds and its relationship to climate. *Ecology* 51: 365–90.

Kaplan, J.M., Pigliucci, M., and Banta, J.A. 2015. Gould on Morton, Redux: What can the debate reveal about the limits of data? *Studies in History and Philosophy of Biological and Biomedical Sciences* 52: 22–31.

Katz, D.C., Grote, M.N., and Weaver, T.D. 2016. A mixed model for the relationship between climate and human cranial form. *American Journal of Physical Anthropology* 160: 593–603.

Legendre, P., and Legendre, L. 2012. *Numerical Ecology*. 3rd English Edition. Amsterdam: Elsevier.

Lewis, J.E., DeGusta, D., Meyer, M.R., Monge, J.M., Mann, A.E., and Holloway, R.L. 2011. The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skulls and bias. *PLoS Biology* 9: e1001071.

Lieberman, L. 2001. How “Caucasoids” got such big crania and why they shrank: from Morton to Rushton. *Current Anthropology* 42: 69–95.

Lindsey, C.C. 1966. Body sizes of poikilotherm vertebrates at different latitudes. *Evolution* 20: 456–65.

Long, J.C., and Kittles, R.A. 2003. Human genetic diversity and the nonexistence of biological races. *Human Biology* 81: 777–98.

Mace, G.M., Harvey, P.H., and Clutton-Brock, T.H. 1981. Brain size and ecology in small mammals. *Journal of Zoology* 193: 333–54.

Manica, A., Amos, W., Balloux, F., and Hanihara, T. 2007. The effect of ancient population bottlenecks on human phenotypic variation. *Nature* 448: 346–48.

Manica, A., Prugnolle, F., and Balloux, F. 2005. Geography is a better determinant of human genetic differentiation than ethnicity. *Human Genetics* 118: 366–71.

Mayr, E. 1942. *Systematics and the Origin of Species*. New York: Columbia University Press.

Mayr, E. 1956. Geographical character gradients and climatic adaptation. *Evolution* 10: 105–8.

Meiri, S. 2011. Bergmann’s Rule – what’s in a name? *Global Ecology and Biogeography* 20: 203–7.

Meiri, S., and Dayan, T. 2003. On the validity of Bergmann’s rule. *Journal of Biogeography* 30: 331–51.

Michael, J.S. 1988. A new look at Morton’s craniological research. *Current Anthropology* 29: 349–54.

Mitchell, P.W. 2018. The fault in his seeds: Lost notes to the case of bias in Samuel George Morton’s cranial race science. *PLoS Biology* 16: e2007008.

Mitchell, P.W., and Michael, J.S. 2019. Bias, brains, and skulls: Tracing the legacy of scientific racism in the nineteenth century works of Samuel George Morton and Friedrich Tiedemann. In August, E., Furrow, B,R., Richter, K., Thomason, K.K., Costello, D., Michael, J.S., Mitchell, P.W., and Bettray, U., editors. *Embodied Difference: Divergent Bodies in Public Discourse* London: Lexington Books.

Montagu, M.F.A. 1964. *The Concept of Race*. New York: Free Press of Glencoe.

Mori, E., Ancillotto, L., Lovari, S., Russo, D., Nerva, L., Mohamed, W.F., Motro, Y., Di Bari, P., and Plebani, M. 2019. Skull shape and Bergmann’s rule in mammals: hints from Old World porcupines. *Journal of Zoology* 308: 47–55.

Morton, S.G. 1839. *Crania Americana; or, a Comparative View of the Skulls of Various Aboriginal Nations of North and South America: To Which Is Prefixed an Essay on the Varieties of the Human Species*. Philadelphia: J. Dobson.

Morton, S.G. 1849. *Catalogue of Skulls of Man and the Inferior Animals, in the Collection of Samuel George Morton*. Philadelphia: Merrihew & Thompson.

Novembre, J. 2022. The background and legacy of Lewontin’s apportionment of human genetic diversity. *Philosophical Transactions of the Royal Society B* 377: 20200406.

Ocobock, C. 2023. Human cold adaptation: an unfinished agenda v2. 0. *American Journal of Human Biology* 36: e23937.

Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., et al., 2022. vegan: Community Ecology Package for R. Accessed 1 April 2023. link to the article.

Osada, N. 2015. Genetic diversity in humans and non-human primates and its evolutionary consequences. *Genes & Genetic Systems* 90: 133–45.

Ousley, S., Jantz, R., and Freid, D. 2009. Understanding race and human variation: why forensic anthropologists are good at identifying race. *American Journal of Physical Anthropology* 139: 68–76.

Pomeroy, E., Stock, J.T., and Wells, J.C.K. 2021. Population history and ecology, in addition to climate, influence human stature and body proportions. *Scientific Reports* 11: 274.

Reich, D. 2018. *Who We are and how We Got Here: Ancient DNA and the New Science of the Human Past*. Oxford: Oxford University Press.

Relethford, J.H. 2004. Boas and beyond: migration and craniometric variation. *American Journal of Human Biology* 16: 379–86.

Relethford, J.H. 2010. Population-specific deviations of global human craniometric variation from a neutral model. *American Journal of Physical Anthropology* 142: 105–11.

Rensch, B. 1924. Das Dépérétsche Gesetz. *Zeitschrift für Indukt. Abstammungs-und Vererbungslehre* 35: 139–55.

Rensch, B. 1938. Some problems of geographical variation and species-formation. In *Proceedings of the Linnean Society of London*. Oxford: Blackwell Publishing Ltd., Vol. 150, pp. 275–85

Richards, R.J. 2018. The beautiful skulls of Schiller and the Georgian girl: quantitative and aesthetic scaling of the races. In Rupke, N., and Lauer, G., editors. *Johann Friedrich Blumenbach: Race and Natural History*. London: Routledge, 1750–50

Ritchie, S.J., Booth, T., Hernández, M.C.V., Corley, J., Maniega, S.M., Gow, A.J., Royle, N.A., Pattie, A., Karama, S., Starr, J.M., et al., 2015. Beyond a bigger brain: Multivariable structural brain imaging and intelligence. *Intelligence* 51: 47–56.

Roberts, D.F. 1954. Body weight, race and climate. *American Journal of Physical Anthropology* 11: 533–58.

Roseman, C.C. 2016. Random genetic drift, natural selection, and noise in human cranial evolution. *American Journal of Physical Anthropology* 160: 582–92.

Roseman, C.C. 2021. Lewontin did not commit Lewontin's fallacy, his critics do: why racial taxonomy is not useful for the scientific study of human variation. *BioEssays* 43: 2100204.

Sánchez-Villagra, M.R. 2022. *The Process of Animal Domestication*. Princeton: Princeton University Press.

Sankararaman, S., Mallick, S., Patterson, N., and Reich, D. 2016. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology* 26: 1241–47.

Savell, K.R.R., Katz, D.C., Weaver, T.D., and Auerbach, B.M. 2022. Mixed models for the relationship between latitude, temperature, and human postcranial form. *American Journal of Biological Anthropology* 179: 431–43.

Schlebusch, C.M., Sjödin, P., Breton, G., Günther, T., Naidoo, T., Hollfelder, N., Sjöstrand, A.E., Xu, J., Gattepaille, L.M., Vicente, M., et al., 2020. Khoe-San genomes reveal unique variation and confirm the deepest population divergence in *Homo sapiens*. *Molecular Biology and Evolution* 37: 2944–54.

Sebro, R., Peloso, G.M., Dupuis, J., and Risch, N.J. 2017. Structured mating: patterns and implications. *PLoS Genetics* 13: e1006655.

Skullerud, K. 1985. Variations in the size of the human brain. *Acta Eurologica Scandinavica. Supplementum* 102: 1–94.

Sokal, R.R., and Rohlf, F.J. 1995. *Biometry*. 3rd Ed. New York: Freeman & Co.

Stringer, C. 2014. Why we are not all multiregionalists now. *Trends in Ecology & Evolution* 29: 248–51.

Stulp, G., and Barrett, L. 2016. Evolutionary perspectives on human height variation. *Biological Reviews* 91: 206–34.

Templeton, A.R. 1998. Human races: a genetic and evolutionary perspective. *American Anthropologist* 100: 632–50.

Templeton, A.R. 2013. Biological races in humans. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 44: 262–71.

Thomson, A. 1913. The correlations of isotherms with variations in the nasal index. In Proceedings of the Seventeenth Congress of Medicine, London, UK, 6–12 August 1913; pp. 89–90.

Tiedemann, F. 1837. *Das Hirn des Negers mit dem des Europäers und Orang-Outangs verglichen*. Heidelberg: K. Winter.

Tobias, P.V. 1970. Brain-size, grey matter and race – fact or fiction? *American Journal of Physical Anthropology* 32: 3–25.

Van Valen, L. 1974. Brain size and intelligence in man. *American Journal of Physical Anthropology* 40: 417–23.

Wagner, J.K., Yu, J.-H., Ifekwunigwe, J.O., Harrell, T.M., Bamshad, M.J., and Royal, C.D. 2017. Anthropologists' views on race, ancestry, and genetics. *American Journal of Physical Anthropology* 162: 318–27.

Watt, C., Mitchell, S., and Sale, V. 2010. Bergmann's rule; a concept cluster? *Oikos* 119: 89–100.

Weisberg, M. 2014. Remeasuring man. *Evolution and Development* 16: 166–78.

Weisberg, M., and Paul, D.B. 2016. Morton, Gould, and Bias: a comment on “the Mismeasure of science”. *PLoS Biology* 14: e1002444.

Whewell, W. 1837. *History of the Inductive Sciences, from the Earliest to the Present Times: In Three Volumes*. London: Parker.

Williams, K.M., Hysi, P.G., Yonova-Doing, E., Mahroo, O.A., Snieder, H., and Hammond, C.J. 2017. Phenotypic and genotypic correlation between myopia and intelligence. *Scientific Reports* 7: 1–8.

Witelson, S.F., Kig, D.L., and Harvey, T. 1999. The exceptional brain of Albert Einstein. *Lancet* 353: 2149–53.

Yong, R., Ranjitkar, S., Lekkas, D., Halazonetis, D., Evans, A., Brook, A., and Townsend, G. 2018. Three-dimensional (3D) geometric morphometric analysis of human premolars to assess sexual dimorphism and biological ancestry in Australian populations. *American Journal of Physical Anthropology* 166: 373–85.