"Transmuting" Women into Men: Galton's Family Data on Human Stature

James A. HANLEY

The first two regression lines, and the first correlations, were calculated by Francis Galton, in his work on heredity in sweet peas and in humans. When "regressing" the heights of adult children on those of their parents, Galton had to deal with the fact that men are generally taller than women—but without modern-day statistical tools such as multiple regression and partial correlation. This article uses the family data on stature, which we obtained directly from Galton's notebooks, to (a) compare the sharpness of his methods, relative to modern-day ones, for dealing with this complication; and (b) estimate the additional familial component of variance in stature beyond that contributed by the parental heights. In keeping with Galton's plea for "a manuscript library of original data," these historical and pedagogically valuable data are now available to the statistical community as digital photographs and as a dataset ready for further analyses.

KEY WORDS: Correlation; Data repository; Historical data; Random-effects model; Regression; Transformations.

1. INTRODUCTION

Francis Galton coined the term *regression* to describe situations in which there is reversion of a characteristic measured in offspring, *away* from the mean value of the *same* characteristic in *their own* parents, and *towards* the mean value in *all* offspring. To measure the "Rate of regression in hereditary stature," Galton (1886) divided the offspring into nine subgroups according to the average height of their two parents (he called this amalgam the "mid-parent"). He plotted the median offspring height against the median mid-parent height, and by eye, fitted a straight line to the nine datapoints. He estimated that "the Deviates of the Children are to those of their Mid-Parents as 2 to 3" implying that "When Mid-Parents are taller than mediocrity, their Children tend to be shorter than they," and conversely.

James A. Hanley is Professor, Department of Epidemiology and Biostatistics, McGill University, 1020 Pine Avenue West, Montreal, Quebec, Canada H3A 1A2 (E-mail: James.Hanley@McGill.CA). I wish to thank Stephen Stigler for his book, and for directing me to the Galton Papers; Rebecca Fuhrer, formerly a professor at University College London (UCL), and now chair of my department at McGill, for introducing me to her UCL graduate student Beverly Shipley; Beverly for locating, extracting, and computerizing the raw data, and urging me to "come to London and see how perfectly Galton had arranged them"; Kate Lewis and archivist Gill Furlong of the Special Collections section of UCL Library Services for encouraging, and arranging for, me to "bring Galton to a wider audience"; Colin Hanley and Louise Koo for photographing the material; Marie-Eve Beauchamp, Carine Bellera, Katia Charland, and Elizabeth Turner for comments on the manuscript; and Editor Fritz Scheuren for his enthusiasm. This work was supported by an operating grant from the Natural Sciences and Engineering Research Council of Canada. The photograph (Figure 1) is published with the permission of the Director of Library Services of University College London.

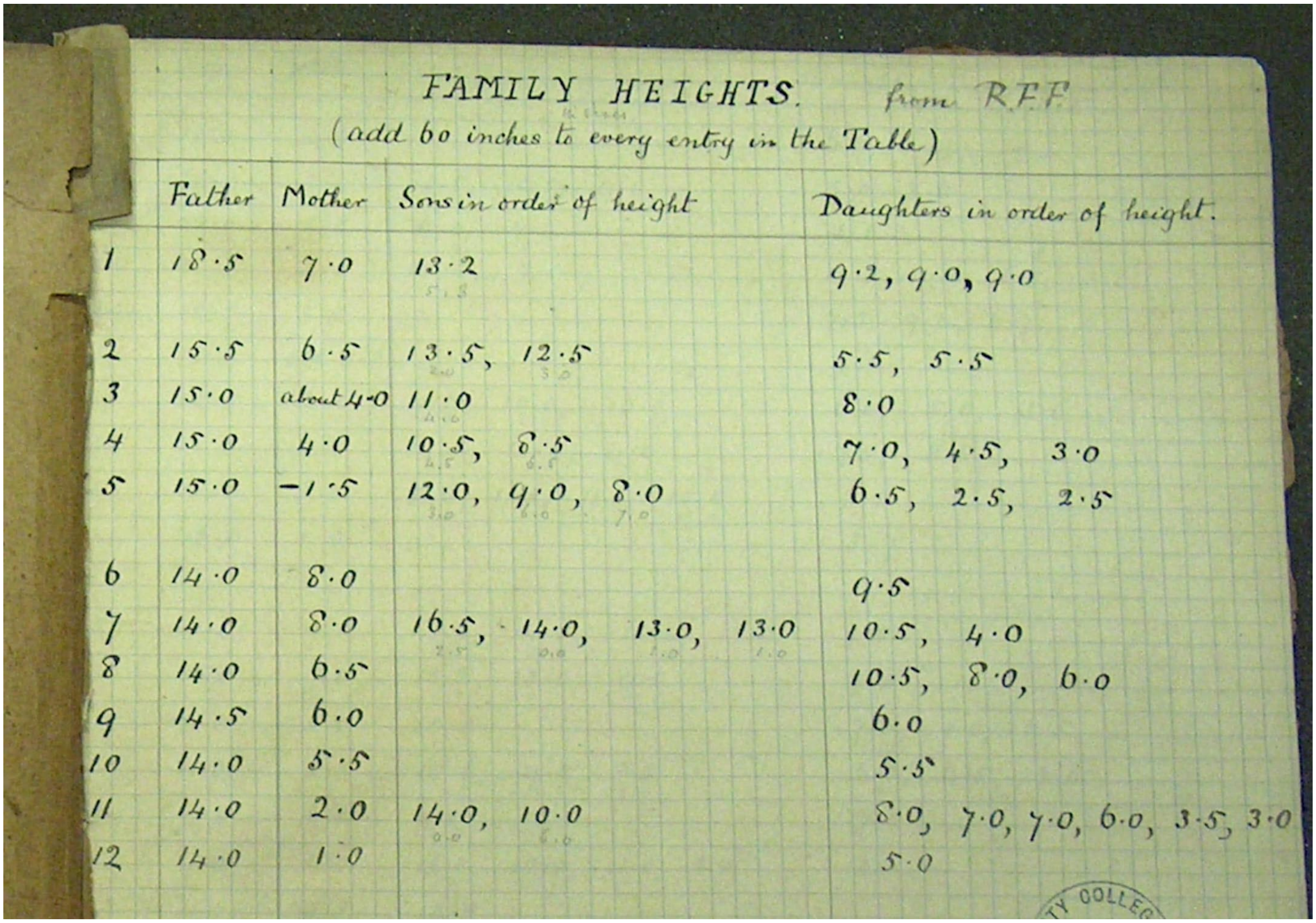
Galton's two-way frequency table "Number of Adult Children of various Statures born of 205 Mid-Parents of various Statures" has a special place in the history of mathematical statistics. The concentric elliptical shape of the contours of equal frequency led Galton to the correlation coefficient of the bivariate Gaussian distribution. From these, Karl Pearson developed a full treatment of correlation, multiple and partial. When illustrating this, Pearson (1896) had not yet accumulated enough data from his own Family Record Series (Pearson and Lee 1903). He therefore relied on "the family data on which (Galton's) work on 'Natural Inheritance' was based." These, "Mr. Galton, with his accustomed generosity, has placed at my disposal."

I became familiar with some of this history after I too was able to obtain these same raw data which Galton placed at Pearson's disposal. This article explains the questions that led me to pursue them, and the answers the data provide. In keeping with Galton's own wishes (1901), I am making "digital hard copies" and electronic versions of these raw data available. Some will merely admire the raw data and how Galton organized them; "others who desire to verify his work" (e.g., Wachsmuth, Wilkinson, and Dallal 2003) can now subject them to statistical analyses that were not possible in 1886, or could not be carried out directly using only the two-way frequency data available until now.

2. BACKGROUND AND QUESTIONS

When teaching regression and correlation, I show students the aggregated data in Table 8.1, and Figures 8.7 and 8.8 from Stigler's (1986) book, but hide how Galton dealt with *the fact that men are generally taller than women*. I ask how they would deal with this "complication" if using the *raw* data *today* to calculate the correlation between the height of the offspring and the mid-parent.

It is easy to imagine a scatterplot of the heights of the 928 offspring in relation to their 928 "mid-parents," and to visualize the two data clouds—the one for sons lying several inches above the one for daughters—that this would produce. Students quickly realize that the single correlation coefficient calculated from the ensemble is attenuated relative to the two separate correlations for sons and for daughters. So, they suggest "partialing out" the "effect" of sex; or "putting sex in the model"; or "adjusting for sex." When challenged as to how they would explain it to a journalist, they begin to see that "adjustment for sex" is conceptually like adding so many inches to the height of each female, or subtracting this amount for each male. Only a few students have suggested the adjustment used by Galton. He tells us that "All female heights were *multiplied by 1.08*"; that is, that he "transmuted" them (1886, p. 247). Even though we did not have the raw data to verify it, students and I usually agree that this "proportional" scaling is a more elegant and biologically appropriate adjustment than the additive one. But the empiricist in



FAMILY HEIGHTS. from RFF

(add 60 inches to every entry in the Table)

	Father	Mother	Sons in order of height	Daughters in order of height.
1	18.5	7.0	13.2	9.2, 9.0, 9.0
2	15.5	6.5	13.5, 12.5	5.5, 5.5
3	15.0	about 4.0	11.0	8.0
4	15.0	4.0	10.5, 8.5	7.0, 4.5, 3.0
5	15.0	−1.5	12.0, 9.0, 8.0	6.5, 2.5, 2.5
6	14.0	8.0		9.5
7	14.0	8.0	16.5, 14.0, 13.0, 13.0	10.5, 4.0
8	14.0	6.5		10.5, 8.0, 6.0
9	14.5	6.0		6.0
10	14.0	5.5		5.5
11	14.0	2.0	14.0, 10.0	8.0, 7.0, 7.0, 6.0, 3.5, 3.0
12	14.0	1.0		5.0

Figure 1. Photograph of the entries for the first 12 families listed in Galton's notebook. Published with the permission of the Director of Library Services of University College London.

me nevertheless wondered: *would today's additive model have been less sharp than Galton's multiplicative one?*

And so, in 2000, I began my search for Galton's "*untransmuted*" data, to determine whether modern-day data analysts, despite stronger computers and user-friendly statistical procedures, would find weaker correlations with the (default) *additive* model than Galton did with his proportional one.

Galton's two-way frequency table and smoothed frequency plot did not identify which children—whom he had already made "unisex"—with the same mid-parental height belonged to which families. Thus, I also hoped that the children would still be found with their families, that is, *before* they were marshalled into what Galton called "filial" arrays. I wished to investigate two additional questions: (i) *among children with the same mid-parental height, do their deviates from the regression line segregate further along family lines, and how might we show this familial variation graphically? (ii) how much would it matter if parental heights were treated as a family-level (i.e., second level) variable in a multilevel analysis?*

3. RAW DATA

Stephen Stigler directed me to the Galton Papers at University College London (UCL). Beverly Shipley, a post-graduate student at UCL, located the material in March 2001. The data were exactly what I had wished, in a single notebook, family by family, with sons and daughters identified, and with all female heights untransmuted. Because of the frail condition of the notebook, photocopying was not permitted, and so she first transcribed the heights of parents and children onto paper, and later from there into a spreadsheet. In February 2003, I requested and obtained permission to digitally photograph the material. We have used these "hard copies" to double-check our data.

Galton obtained the data "through the offer of prizes" for the "best Extracts from their own Family Records" (Galton 1889, p. 72). Each family correspondent used an album designed by Galton. The author's Web site (Hanley 2004) shows photographs of the cover, and of some pages from, the one completed "Record of Family Faculties" (RFF) extant.

Figure 1 is a photograph of the top half of the first page of the notebook entitled, in Galton's handwriting, "Copies of original data: RFF and special returns of brothers." The figure shows the entries for the first 12 families listed. Families are sorted according to the father's, and within these, the mother's height. In keeping with his promise to contributors, Galton had removed all family identifiers. Within each family, sons are listed first, in order of decreasing height, followed by daughters, sorted similarly. All heights are entered as deviations from 60 inches. The regression to mediocrity is evident in these extreme cases.

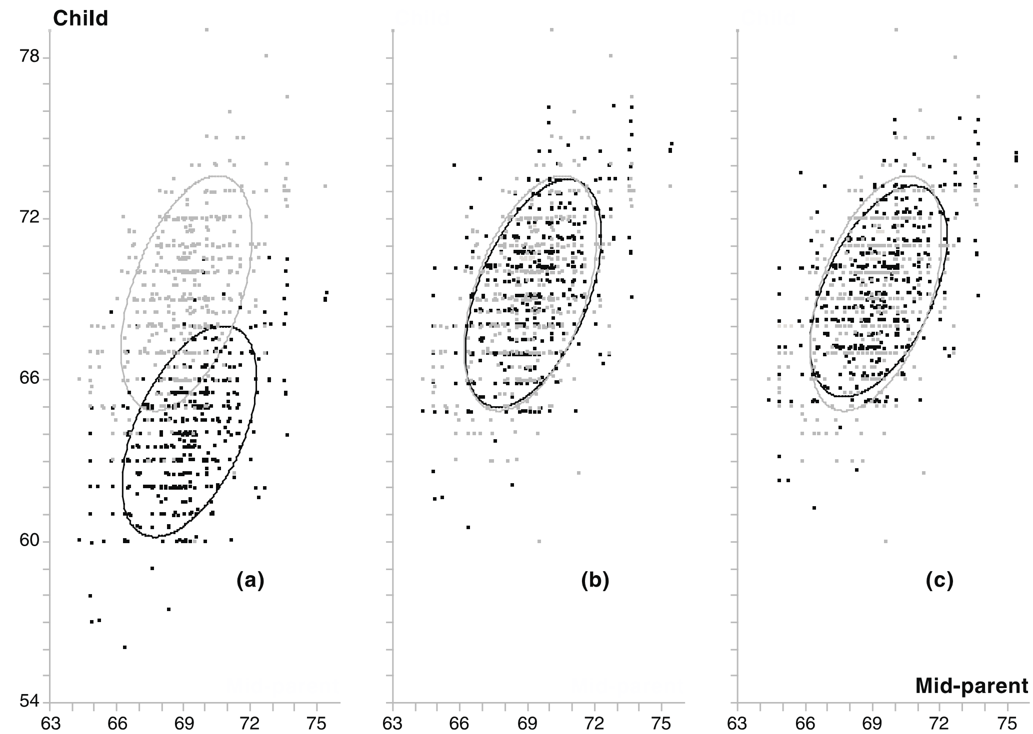


Figure 2. Heights (in inches) of adult children in relation to their mid-parent height. (a) each daughter's height "as is" (b) daughter's height multiplied by 1.08 (c) 5.2 inches added to daughter's height. Daughters' heights are shown in darker, and sons' in lighter, symbols. Ellipses (75%) are drawn based on the observed means and covariances. In all three panels, and in analyses for Figure 3, the mid-parent height is calculated as (father's height + 1.08 × mother's height)/2.

The entire "listing" contains entries for 963 children (486 sons, 476 daughters) from 205 families ranging in size from 1 to 15 children. Some 934 children had numerical values (35 were recorded as "*about* x.0 inches"). In 26 others (21 female, 5 male) height was described verbally ("tallish," "middle," etc.); two individuals were noted as "deformed" and one other as "idiotic." Although Galton (1886, p. 247; 1889, p. 77) referred to the heights of 9<u>30</u> adult children, his table shows—and on page 91 in his 1889 text he speaks of—9<u>28</u> adult offspring. Because I was unable—from his frequency distribution—to confidently identify these 928 from among the 934 we found in the "Copies of original data," I will compare the various methods of analysis using the 934.

4. ANALYSES

Multiplicative or additive: which model is sharper? Galton used medians rather than means, and an ad-hoc method of fitting a line to the nine datapoints. I used least-squares regression in order to narrow our now-versus-then comparison to how Galton and we might scale the daughters' heights, that is, to the sharpness of Galton's multiplicative, versus the default modern-day additive, scaling.

The raw data, and the unisex data, with female heights transmuted both by Galton's transformation of both their mean and variance, or just their mean, are shown in Figure 2. Contrary to what Galton did in his Table and in Plate X, I show children's heights on the vertical, and the mid-parent height on the horizontal axis. The way he sorted and organized the data in his notebook—with each family in a different row, and the children spread out along the row—may explain why he oriented his frequency table, and the figure drawn from it, the way he did. However, even though the diagram in Plate IX does not have an *explicitly marked* horizontal axis (he uses the vertical axis for both parents and children), it is clear from the nine datapoints and the line of identity on this diagram that children's heights are on the vertical, and parental heights on the horizontal, axis—just as we would orient them today.

Table 1 shows the correlations and regression coefficients obtained by the various analysis methods. Rows 1 and 2 show, as expected, that the *correlation* is greatly attenuated if one does not take account of the sex of the offspring. The *slope* is largely unaffected—the fact that it is lower in the ensemble (0.64) than in each sex separately (0.66 and 0.71) can be partly explained by the fact that the average mid-parent height is 69.3 inches in daughters, greater than the corresponding average of 69.1 inches for the sons, a sizable difference given the already relatively narrow range of mid-parent heights.

Rows 3 and 4 contrast the effect of *multiplying* daughters' heights by 1.08 and the modern-day blackbox approach which, in effect, *adds* 5.2 inches to each daughter's height. Galton's external, "low-tech," and biologically attractive scaling gives a higher slope, and—as seen in Figures 2(b) and 2(c)—more similar sex-specific ellipses. But the difference in slopes is quite small (0.71 versus 0.69). The correlations obtained by the multiplicative and additive methods are identical to two decimal places (0.50), and the RMSE's agree to 1 decimal place (2.2 inches).

The R^2 of 0.63 for the analysis based on the multiple regression strategy does not mean that this *model* is better than those that produce R^2's of 0.23 to 0.26. Rather, the "improvement" is because the *data* are different: by design, heights are

Table 1. *Regression Towards Mediocrity: Results of Different Modern-Day Strategies for Dealing With the Fact that Sons are Generally Taller than Daughters*. Children's adult heights are regressed on/correlated with mid-parent height (i.e., average of father's and transmuted mother's height).*

Strategy	*r*	*slope (b)*	*RMSE**	R^2
Simple correlation/linear regression, without regard to sex	0.32	0.64	3.4"	0.10
Sex-specific**: simple correlation/linear regression				
Daughters	0.51	0.66	2.0"	0.26
Sons	0.48	0.71	2.3"	0.23
Simple correlation/linear regression, after daughters' heights have been multiplied by 1.08	0.50	0.71	2.2"	0.25
Partial correlation/multiple linear regression***; "untransmuted" heights	0.50	0.69	2.2"	0.63

* RMSE: Root mean squared error.

** Daughters mean(SD): 64.1(2.4)″, *n* = 453. Sons: 69.2(2.6)″, *n* = 481; The corresponding values for their mid-parent heights are 69.3(1.8)″ and 69.1(1.8)″.

*** 64.1″ + 5.2″ × (male?) + 0.69 × (midparental deviation from mediocrity of 69.2″)

already more homogeneous (single sex or unisex) in these latter instances. In contrast, the analysis that yields the R^2 of 0.63 is performed on the raw mixed-sexes data and thus "takes credit" for "explaining" the heterogeneity in these raw data.

"*Galton's bend?*" Nonlinear regressions, quadratic and cubic, fitted to the data in Figure 2 did not significantly improve the fit over a linear one. With mid-parent height (h_p) centered at $c = 69.2$ inches, the quadratic equation was $0.701(h_p - c) + 0.030(h_p - c)^2$ for "transmuted" daughters $(\mathrm{SE}[0.030] = 0.019)$ and $0.713(h_p - c) + 0.014(h_p - c)^2$ for sons $(\mathrm{SE}[0.014] = 0.021)$. The coefficient of the quadratic term was 0.022 $(\mathrm{SE} : 0.014, P = 0.11)$ when a single equation was fitted to the pooled data. This curvature is a good deal less than what Wachsmuth et al. (2003) found using the originally tabulated frequencies for the pooled data; they applied their smoother to the *inverse* regression, that of *mid-parent on child*. Interestingly, I also found more curvature when I applied a smoother to the inverse regression.

Even if it is not clear what *criterion* Galton used to fit his line, there are also differences in the *data* used in the 1886 and 2003 analyses. Galton fitted his straight line (in Plate IX, 1886) using *only the medians* in *only nine* filial arrays (he included in the same "filial array" all the offspring whose mid-parent height was in the same one-inch interval). In contrast, the 2003 analysis used the midpoint of, and associated frequency for, each of the $1'' \times 1''$ bins for each of *11* filial arrays; it was not clear what midpoints the 2003 analysis used for the 18 open-ended bins, containing 36 offspring, at the edges of the frequency table. In a footnote to the identical table, reprinted in his 1889 book, Galton tells us that the topmost of the 11 filial arrays (i.e., the top row in his 11-row frequency table) with four children from five families—an error which he acknowledged, but did not correct—was "not considered at all" because of "the paucity of the numbers it contains." Nor does he seem to have used the bottom filial array, although in the same 1889 footnote, he does assure us that this array, with 14 children in 1 family, "which looks suspicious, is correct." Further evidence that he did not use the families with the very tallest and very shortest mid-parents comes from the fact that Galton's iso-frequency contour for the bivariate distribution (Plate X, 1886) is overlaid onto *only seven* rows of smoothed frequencies, derived, he tells us, by taking two-dimensional moving averages of the frequencies in the nine filial arrays.

Incidentally, Galton's notebook lists the heights of *15* (rather than 14) adult children for this one (very short mid-parent, 64.9 inches) family: the 15 are shown in a column at the left side of the scatterplots in Figure 2. To the immediate left of this large (but short mid-parent) family are the two offspring from a family with an even shorter mid-parent (64.4 inches); we were unable to identify these two offspring in Galton's frequency table. It is not clear whether these and the remaining discrepancies (934 observations in the notebook versus 928 in the table) were simple data handling errors, or deliberate data selectivity. Nor is it obvious whether the "listing" shown in Figure 1 was compiled before or after the data-analysis. If, as I suspect, all of the data analyses were carried out from the frequency table, and the listing was compiled only afterwards, then the discrepancies are easily imagined. One must also wonder, in the case of Pearson and Lee's (1903) 78 two-way frequency tables, whether they ever compiled a complete "data listing," or whether the frequencies were simply tallied directly from the original data forms.

Do the deviates from the regression line segregate further along family lines? Galton fit his regression line using just nine datapoints, each one corresponding to a "filial array," that is, all offspring with a mid-parent height in the same one-inch interval. The shortest mid-parent category used was 64–65 inches and the tallest was 72–73. The numbers of offspring in the 9 arrays used ranged in size from 23 offspring of the 5 families with the shortest mid-parents, to 219 offspring of 49 families with mediocre mid-parents, to 19 offspring of 6 families with the tallest mid-parents. (The separate weighted and unweighted regression analyses I applied to the nine data points in Galton's Plate IX suggest that when fitting his line "by straight edge" (Pearson 1930), Galton gave equal weight to each of the nine medians).

Knowing which offspring are from which families, it is now possible to estimate to what extent the deviates of the 934 offspring from the simple regression line segregate further along family lines.

To represent and estimate the extent of the further segregation by family, one might model the height of the jth offspring of family i so that the deviation from the overall regression line is partitioned into independent between- and within-family components. One can write this random effects model as

$$\text{height}_{i,j} = \mu^* + \beta^* \times \text{height of mid-parent}_i + \alpha_i + \varepsilon^*_{i,j}, \quad (1)$$

where α_1 to α_{205} are the between-family components, and the $\varepsilon^*_{i,j}$ are the n_i within-family ones. Modern software can simultaneously estimate the standard deviations σ_B and σ_W of these two (between- and within-family) components. The estimates of σ^2_B and σ^2_W were 0.96 in.^{2} and 4.11 in.^{2} when fitted using the MIXED procedure in SAS (1996), and virtually the same when fitted by WinBUGS (Spiegelhalter, Thomas, and Best 1999). The resulting intraclass correlation of $0.96/(0.96 + 4.11) = 0.19$ agrees with the pairwise (exchangeable) within-family correlation of 0.19 estimated by a GEE approach.

Alternatively, taking it in two, albeit slightly less sophisticated steps, one could first fit, and extract the residuals from, (Galton's) model

$$\text{height}_{i,j} = \mu + \beta \times \text{height of mid-parent}_i + \varepsilon_{i,j} \quad (2)$$

and then subject the calculated $e_{i,j}$'s to a one-way components of variance analysis. But how to show this familial component in a graphical way, using the *original data* in a simple dot plot that amplifies the quite modest ICC and makes it visible to the human eye? Incidentally, when Galton wished to show that there was virtually no correlation between heights of fathers and mothers (but had not yet developed an index of correlation), he did so indirectly by showing that the variance of the sum of the parental heights was only slightly greater than the sum of their variances!

If one computed the mean, $\overline{e}_i$, for family i, then *even if σ_B were zero*, the amplitude of the series of means $\{\overline{e}_1, \ldots, \overline{e}_{205}\}$ would vary across i with the number of family members contributing to each mean: σ_W in single-offspring families, half this amount in four-offspring families, and so on. Thus, to ensure that the 205 family-level deviates were all on the same scale, I scaled $\overline{e}_i$ upwards by $n_i^{1/2}$.

Transformations were also required when extracting and visually displaying the within-family variation in the e's, in order to ensure that (i) the scale was common both within and across different-size families and (ii) within family i, only $n_i - 1$ uncorrelated deviations from $\overline{e}_i$ were selected from the n_i correlated ones. For example, suppose we knew that the deviation of a selected child was greater than the average of all three within-family deviations, that is, those of the selected child and its two siblings; this contains information about the deviations of the other two siblings. Therefore, if family i had two or more offspring, I converted the n_i residuals into $n_i - 1$ orthogonal differences. To do so, I first randomly ordered the n_i family members into a list, which for illustration will be indexed by k. For the kth family member in this ordered list, $[1 \leq k \leq (n_i - 1)]$, I then computed the difference between the residual of this kth individual and the mean of the $(n_i - k)$ residuals of those family members *later* in the list, to form the *orthogonal* differences

$$d_{i,k} = [(n_i - k)/(n_i - k + 1)]^{1/2} \times \{e_{i,k} - \text{mean}(e_{i,k+1} \text{ to } e_{i,n_i})\}.$$

Note that each successive difference (the first minus the average of the second and beyond, the second minus the average of the third and beyond, etc.) was scaled up by a different amount to ensure a common amplitude both within and across different-size families. The 172 families with two or more offspring yielded 729 orthogonal within-family differences. I ignored the fact that the 934 residuals from model (2) have only 932 degrees of freedom, or that the residuals at the extremes of mid-parental height are, by construction, slightly less variable.

The 729 scaled within-family differences are displayed in Figure 3; the 205 scaled means (average family deviations) are overlaid on them. In an informal survey, where I asked students and colleagues to visually judge which series had the greater amplitude, a majority perceived the between-family series to be more variable. The boxplot markers, which I added later, confirm that there is indeed additional *between*-family variation—the variances of the 205 and the 729 computed quantities were 8.33 in.^{2} and 4.11 in.^{2}, respectively. The detectability of this additional component of familial variance was enhanced by the strategy of amplifying the mean of the e's from the same family by the square root of the number of offspring: if n is the (average) number of offspring per family, then the expected square of the between-family deviations is $\sigma^2_W + n \times \sigma^2_B$, rather than the $\sigma^2_W + \sigma^2_B$ that serves as the denominator of the ICC. Had the *squares* of the between-family quantities been ordered by, and plotted against, family size, rather than the order shown, the familial component would have been even more obvious, and σ^2_B estimable from the slope of the empirical relationship between the squares and n.

The 205 average deviations include not just familial variation, but also any lack of fit of the linear regression line. However, when the deviations of the 934 heights in Figure 2(b) were measured from a quadratic regression curve, the estimates of σ^2_W and σ^2_B were still 0.96 in.^{2} and 4.11 in.^{2}, respectively—differing only in the third decimal place from than those calculated from the straight line model.

5. DISCUSSION

As can be seen by comparing the overlaid ellipses in Figures 2(b) and 2(c), scaling daughters' heights multiplicatively, so that their variance is also increased, does fit slightly better than a simple additive shift. However, the "adjusted" correlation and regression coefficient obtained from the two models were quite similar in magnitude. In retrospect, this is probably not that surprising. Transmuting a daughter who is 4′ 8″—the shortest of the 453—using Galton's model adds 4.5 inches to her height, while the "same correction for all" model used in the default modern-day analysis adds 5.2 inches. At the other extreme, the 5′10.5″ daughter—the tallest in the dataset—is transmuted by 5.6 inches. Most discrepancies between the two methods are much smaller than these, and thus quite small in relation to the large range of heights. Given this, and the fact that the quality of the reported heights was not completely standardized, the lim-

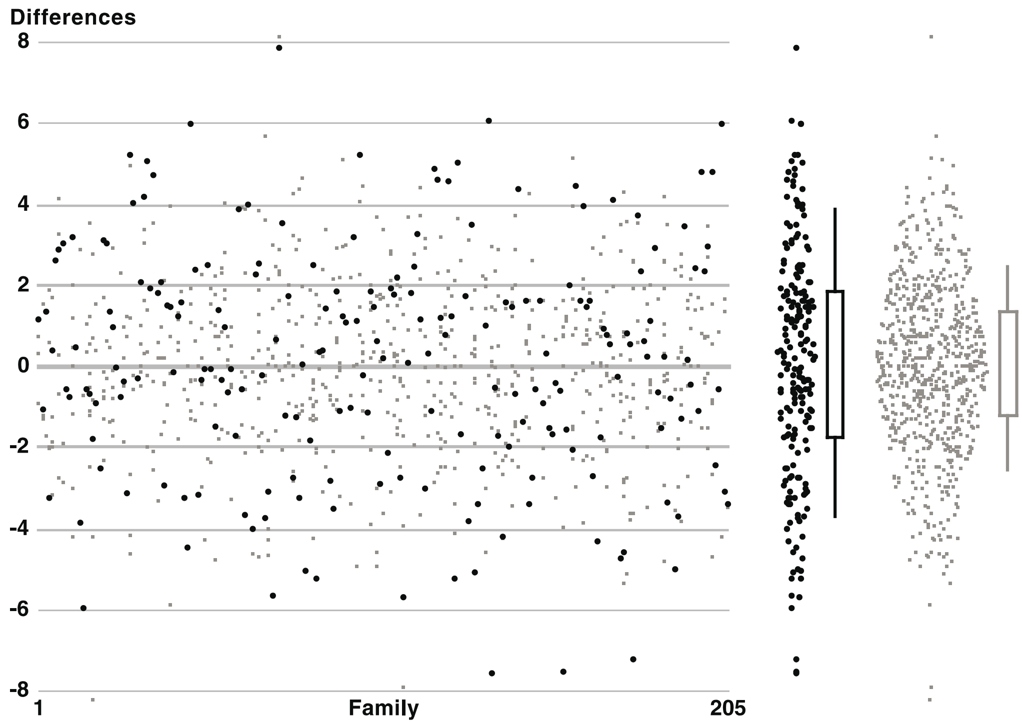


Figure 3. Distribution of within- and between-family residuals from simple linear regression, after daughters' heights have been multiplied by 1.08, of offspring height on mid-parent height. Families listed left to right, in same order as in Galton's notebook. Larger darker dot: the average residual for a family, multiplied by the square root of the number of offspring in the family, so as to put all 205 averages on the same scale. Smaller lighter dot: orthogonal difference of within-family residuals (729 in all, from 172 families with two or more offspring; see text). Marginal distributions shown on right. Boxplots show the 10th, 25th, 75th, and 90th percentiles.

ited impact of the choice of method of scaling is understandable. [Pearson also surmised, from the weaker parent-child correlations in females than in males (1896, Table II, p. 270; 1930, p. 18), that "it may well have been due to amateur measuring of stature in women" (1930, p. 18).]

Pearson (1930, p. 8 and again on p. 15) discussed—and on theoretical grounds seemed to favor—yet another method of transmuting heights. He wondered why Galton, who suggested it much earlier (1877, p. 283), changed methods later when dealing with human stature. In his earlier method, Galton suggested: "Suppose, for example, a female whose height was equal to the average female height $+ 3^\circ$ of female deviation, the equivalent in terms of male stature is the average male height $+ 3^\circ$ of male deviation. Hence the female in question must be registered not in feet and inches of her actual height, but in those of the equivalent male stature." (Galton actually worked with the median rather than the mean, and with "Q" = half the interquartile range—his measure of the "probable error"—rather than standard deviation. Thus, a 3° deviation corresponds to approximately $3 \times 0.67 =$ 2 standard deviations). Applying this method to the daughters' heights before merging them with the sons' heights yields a correlation coefficient of 0.50, and a coefficient of regression of 0.72, a tiny improvement on the 0.71 obtained by the use of the 1.08. Pearson (1930, p. 15) noted that "practically, the two methods will only agree, if the ratio of the two variabilities is equal to the ratio of the two means, i.e., if the so-called coefficients of variability (c.v.'s) of the two sexes are equal." In Galton's data the c.v. for daughters is 3.7%; for sons it is 3.8%.

I have used Galton's definition of mid-parent, so as to limit comparisons to methods for dealing with heights of *offspring*. I will not discuss whether/how a *mother's* height should be transmuted/combined with that of the father. Pearson addressed this issue in detail in 1896, and again in 1930. The modern definition of a mid-parent (the simple average of the mother's and father's heights, used in managing children with growth disorders) was re-revisited by Cole (2000), using extensive contemporary data. Possibly because of his work with the modern-day version, Cole was under the impression that while Galton multiplied the *daughters'* heights by 1.08, he did not do the same for *mothers*. But Galton states in his frequency table, and in his bivariate contours diagram, that "*all* female heights were multiplied by 1.08″ (italics mine). The median mid-parent of 68.25″ used in his writings also suggests, and the raw data described here now confirm, that he *also transmuted mothers' heights before averaging the two*. Cole (2000, p. 401) showed theoretically that "(i) the simple average of the two parents' heights does not treat the two parents equally in centile terms; (ii) increasing the mother's height by a constant factor before averaging is a suitable way to compensate." In the several data series he examined, he found that "the optimal value for the factor is close to 1.08"—vindicating what Galton did in fact do. Cole arrived at conclusion (ii) by averaging the two parents' z-scores, a procedure close to that suggested by Galton in 1877.

We leave the final word on the 1.08 to Galton (1889, p. 78), and to his computer (italics mine).

> The factor I used was 1.08, which is equivalent to adding a little less than one-twelfth to each female height. It differs slightly from the factors employed by other anthropologists, who, moreover, differ a trifle

between themselves; anyhow, it suits my data better than 1.07 or 1.09. I can say confidently that the final result is not of a kind to be sensibly affected by these minute details, because it happened that owing to a mistaken direction, the *computer to whom* I first entrusted the figures used a somewhat different factor, yet the results came out closely the same.

One would require additional data before trying to estimate how much of the *familial* σ_B^2 is attributable to nature (preconception), and how much to nurture (post-conception). Despite this, instructors should still find the dataset a useful starting point when introducing multilevel models, and illustrating the effects of using variables, such as mid-parental height, as if they were offspring-level variables. I leave it to readers to predict, or if unsure, to determine empirically, whether the point and interval estimates of β^* and the other parameters in model (1) [or β in model (2)] would be different if (a) in this dataset mid-parental height were treated as an offspring- rather than a family-level variable; (b) there were just 30 offspring, 2 families of size 15, or 6 of size 5, or 15 of size 2; or (c) we had information on birth order, and family size, income, and other circumstances when these offspring were growing up.

6. CONCLUDING REMARKS

Galton, helping Pearson launch *Biometrika* (1901), wrote

(. . .) This journal, it is hoped, will justify its existence by supplying these requirements either directly or indirectly. I hope moreover that some means may be found, through its efforts, of forming a manuscript library of original data. Experience has shown the advantage of occasionally rediscussing statistical conclusions, by starting from the same documents as their author. I have begun to think that no one ought to publish biometric results, without lodging a well arranged and well bound manuscript copy of his data in some place where it should be accessible, under reasonable restrictions, to those who desire to verify his work.

For close to a century, the important historical and pedagogic data described here have resided only in the pages of the now-frail notebook "Copies of original data: RFF" in the Galton Papers, out of reach of, or even unknown to, the many researchers and teachers who might wish to use them (e.g., Wachsmuth et al. 2003; Wilkinson 2003). For all we know, the last person to consult and use them may have been Karl Pearson (I cannot determine who added the penciled-in computations under the sons' heights, or to what purpose). Galton, with his flair for the technological—for online access to his papers and biographies, see Tredoux (2004)—would have welcomed the Internet, "computers" that follow instructions, and digital photography. He would also have been pleased that, with the approval of University College London, digital photographs of the pages of his notebook of heights, along with an electronic copy of the numbers they contain, and some other related photographs, are available at http://www.epi.mcgill.ca/hanley/galton.

[Received May 2004. Revised June 2004.]

REFERENCES

Cole, T. J. (2000), "Galton's Midparent Height Revisited," *Annals of Human Biology*, 27, 401–405.

Galton, F. (1877), "Typical Laws of Heredity," in *Proceedings of the Royal Institution of Great Britain*, 8, pp. 282–301.

——— (1886), "Regression Towards Mediocrity in Hereditary Stature," *Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246–263.

——— (1889), *Natural Inheritance*, London: Macmillan.

——— (1901), "Biometry," *Biometrika*, 1, 7–10.

——— (1908), *Memories of My Life* (2nd ed.) London: Methuen.

Hanley, J. A. (2004), Digital photographs of data in Galton's notebooks, and related material, available online at http://www.epi.mcgill.ca/hanley/galton.

Pearson, K. (1896), "Mathematical Contributions to the Theory of Evolution. III Regression, Heredity and Panmixia," *Philosophical Transactions of the Royal Society of London*, Series A, 187, 253–318.

——— (1930), *The Life, Letters and Labours of Francis Galton*, (Vol. IIIA), London: Cambridge University Press.

Pearson, K., and Lee, A. (1903), "On the Laws of Inheritance in Man: I. Inheritance of Physical Characters," *Biometrika*, 2, 357–462.

SAS Institute (1996), Software Release 6.12, Cary, NC: SAS Institute Inc.

Spiegelhalter, D. J., Thomas, A., and Best, N. G. (1999), *WinBUGS Version 1.2 User Manual*, MRC Biostatistics Unit, Cambridge University.

Stigler, S. (1986), "The English Breakthrough: Galton," in *The History of Statistics: The Measurement of Uncertainty before 1900*, Cambridge, MA: The Belknap Press of Harvard University Press, chap. 8.

Tredoux, G. (2004), Web site http://www.galton.org.

Wachsmuth, A., Wilkinson, L., and Dallal, G. E. (2003), "Galton's Bend: A Previously Undiscovered Nonlinearity in Galton's Family Stature Regression Data," *The American Statistician*, 57, 190–192.

Wilkinson, L. (2003), "Smoothers," http://www.spss.com/research/wilkinson/Applets/smoothers.html.