ANTHROPOLOGICAL MISCELLANEA.

REGRESSION *towards* MEDIOCRITY *in* HEREDITARY STATURE.

By FRANCIS GALTON, F.R.S. &C.

[WITH PLATES IX AND X.]

THIS memoir contains the data upon which the remarks on the Law of Regression were founded, that I made in my Presidential Address to Section II, at Aberdeen. That address, which will appear in due course in the Journal of the British Association, has already been published in "Nature," September 24th. I reproduce here the portion of it which bears upon regression, together with some amplification where brevity has rendered it obscure, and I have added copies of the diagrams suspended at the meeting, without which the letterpress is necessarily difficult to follow. My object is to place beyond doubt the existence of a simple and far-reaching law that governs the hereditary transmission of, I believe, every one of those simple qualities which all possess, though in unequal degrees. I once before ventured to draw attention to this law on far more slender evidence than I now possess.

It is some years since I made an extensive series of experiments on the produce of seeds of different size but of the same species. They yielded results that seemed very noteworthy, and I used them as the basis of a lecture before the Royal Institution on February 9th, 1877. It appeared from these experiments that the offspring did *not* tend to resemble their parent seeds in size, but always to be more mediocre than they—to be smaller than the parents, if the parents were large; to be larger than the parents, if the parents were very small. The point of convergence was considerably below the average size of the seeds contained in the large bagful I bought at a nursery garden, out of which I selected those that were sown, and I had some reason to believe that the size of the seed towards which the produce converged was similar to that of an average seed taken out of beds of self-planted specimens.

The experiments showed further that the mean filial regression towards mediocrity was directly proportional to the parental deviation from it. This curious result was based on so many plantings, conducted for me by friends living in various parts of the country, from Nairn in the north to Cornwall in the south, during one, two, of even three generations of the plants, that I could entertain no doubt of the truth of my conclusions. The exact ratio of regression remained a little doubtful, owing to variable influences; therefore I did not attempt to define it. But as it seems a pity that no record could exist in print of the general averages, I give them, together with a brief statement of the details of the experiment, in Appendix I to the present memoir.

After the lecture had been published, it occurred to me that the grounds of my misgivings might be urged as objections to the general conclusions. I did not think them of moment, but as the inquiry had been surrounded with many small difficulties and matters of detail, it would be scarcely possible to give a brief and yet a full and adequate answer to such objections. Also, I was then blind to what I now perceive to be the simple explanation of the phenomenon, so I thought it better to say no more upon

the subject until I should obtain independent evidence. It was anthropological evidence that I desired, caring only for the seeds as means of throwing light on heredity in man. I tried in vain for a long and weary time to obtain it in sufficient abundance, and my failure was a cogent motive, together with others, in inducing me to make an offer of prizes for Family Records, which was largely responded to, and furnished me last year with what I wanted. I especially guarded myself against making any allusion to this particular inquiry in my prospectus, lest a bias should be given to the returns. I now can scarcely contemplate the possibility of the records of height having been frequently drawn up in a careless fashion, because no amount of unbiassed inaccuracy can account for the results, contrasted in their values but concurrent in their significance, that are derived from comparisons between different groups of the returns.

An analysis of the Records fully confirms and goes far beyond the conclusions I obtained from the seeds. It gives the numerical value of the regression toe ward mediocrity in human stature, as from 1 to $\frac{2}{3}$ with unexpected coherence and precision [*see* Plate IX, fig. (*a*)], and it supplies me with the class of facts I wanted to investigate—the degrees of family likeness in different degrees of kinship, and the steps through which special family peculiarities become merged into the typical characteristics of the race at large.

My data consisted of the heights of 930 adult children and of their respective parentages, 205 in number. In every case I transmuted the female statures to their corresponding male equivalents and used them in their transmuted form, so that no objection grounded on the sexual differences of stature need be raised when I speak of averages. The factor I used was 1.08, which is equivalent to adding a little less than one-twelfth to each female height. It differs a very little 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. The final result is not of a kind to be affected by these minute details, for 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.

I shall now explain with fulness why I chose stature for the subject of inquiry, because the peculiarities and points to be attended to in the investigation will manifest themselves best by doing so. Many of its advantages are obvious enough, such as the case and frequency with which its measurement is made, its practical constancy during thirty-five years of middle life, its small dependence on differences of bringing up, and its inconsiderable influence on the rate of mortality. Other advantages which are not equally obvious are no less great. One of these lies in the fact that stature, is not a simple element, but a sum of the accumulated lengths or thicknesses of more than a hundred bodily parts, each so distinct from the rest as to have earned a name by which it can be specified. The list of them includes about fifty separate bones, situated in the skull, the spine, the pelvis, the two legs, and the two ankles and feet. The bones in both the lower limbs are counted, because it is the average length of these two limbs that contributes to the general stature. The cartilages interposed between the bones, two at each joint, are rather more numerous than the bones themselves. The fleshy parts of the scalp of the head and of the soles of the feet conclude the list. Account should also be taken of the shape and set of many of the bones which conduce to a more or less arched instep, straight back, or high head. I noticed in the skeleton of O'Brien, the Irish giant,

TABLE I.

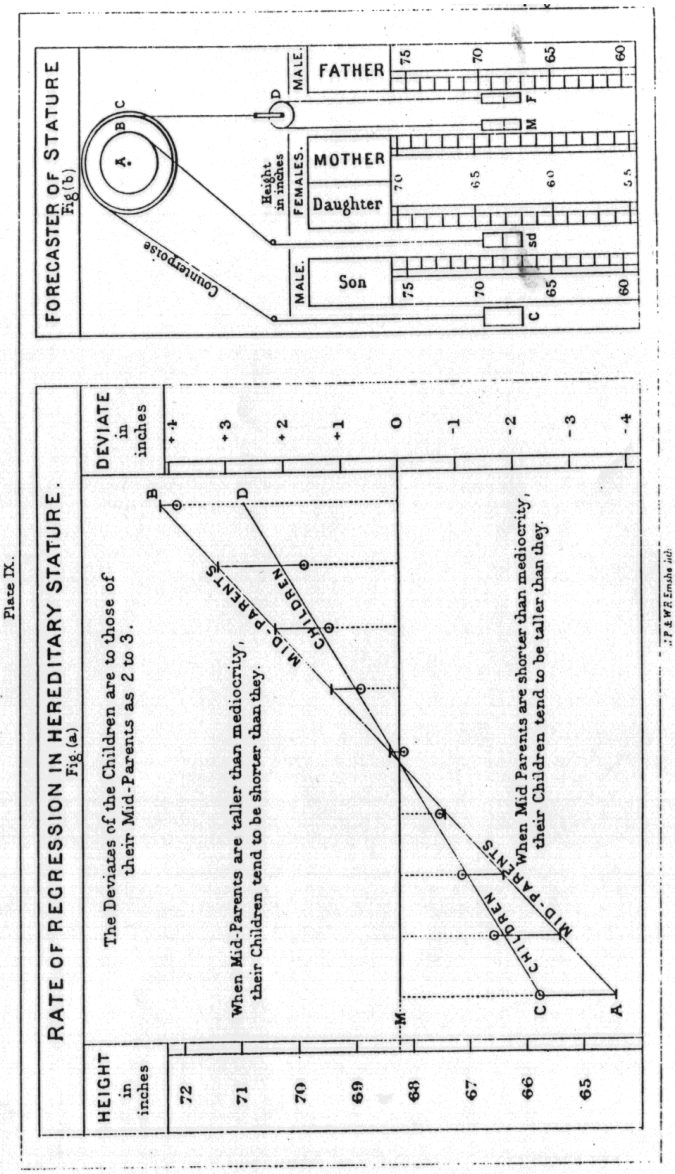
NUMBER OF ADULT CHILDREN OF VARIOUS STATURES BORN OF 205 MID-PARENTS OF VARIOUS STATURES.

(All Female heights have been multiplied by 1·08).

Heights of the Mid-parents in inches.	Heights of the Adult Children.														Total Number of		Medians.
	Below	62·2	63·2	64·2	65·2	66·2	67·2	68·2	69·2	70·2	71·2	72·2	73·2	Above	Adult Children.	Mid-parents.	
Above ..	..	..	..	..	..	..	..	..	..	..	..	1	3	..	4	5	..
72·5	..	..	..	..	..	..	..	1	2	1	2	7	2	4	19	6	72·2
71·5	..	..	..	..	1	3	4	3	5	10	4	9	2	2	43	11	69·9
70·5	1	..	1	..	1	1	3	12	18	14	7	4	3	3	68	22	69·5
69·5	..	..	1	16	4	17	27	20	33	25	20	11	4	5	183	41	68·9
68·5	1	..	7	11	16	25	31	34	48	21	18	4	3	..	219	49	68·2
67·5	..	3	5	14	15	36	38	28	38	19	11	4	..	..	211	33	67·6
66·5	..	3	3	5	2	17	17	14	13	4	..	..	..	..	78	20	67·2
65·5	1	..	9	5	7	11	11	7	7	5	2	1	..	..	66	12	66·7
64·5	1	1	4	4	1	5	5	..	2	..	..	..	..	..	23	5	65·8
Below ..	1	..	2	4	1	2	2	1	1	..	..	..	..	..	14	1	..
Totals ..	5	7	32	59	48	117	138	120	167	99	64	41	17	14	928	205	..
Medians ..	..	..	66·3	67·8	67·9	67·7	67·9	68·3	68·5	69·0	69·0	70·0	..	..	..	..	..

NOTE.—In calculating the Medians, the entries have been taken as referring to the middle of the squares in which they stand. The reason why the headings run 62·2, 63·2, &c., instead of 62·5, 63·5, &c., is that the observations are unequally distributed between 62 and 63, 63 and 64, &c., there being a strong bias in favour of integral inches. After careful consideration, I concluded that the headings, as adopted, best satisfied the conditions. This inequality was not apparent in the case of the Mid-parents.

Plate IX.


RATE OF REGRESSION IN HEREDITARY STATURE.
Fig. (a)
The Deviates of the Children are to those of their Mid-Parents as 2 to 3.
When Mid-Parents are taller than mediocrity, their Children tend to be shorter than they.
When Mid Parents are shorter than mediocrity, their Children tend to be taller than they.
MID-PARENTS
CHILDREN
HEIGHT in inches
DEVIATE in inches
72
71
70
69
68
67
66
65
+4
+3
+2
+1
0
-1
-2
-3
-4
M
A
B
C
D
FORECASTER OF STATURE
Fig. (b)
Counterpoise
Height in inches
MALE.
FEMALES.
Son
Daughter
MOTHER
FATHER
75
70
65
60
C
sd
M
F

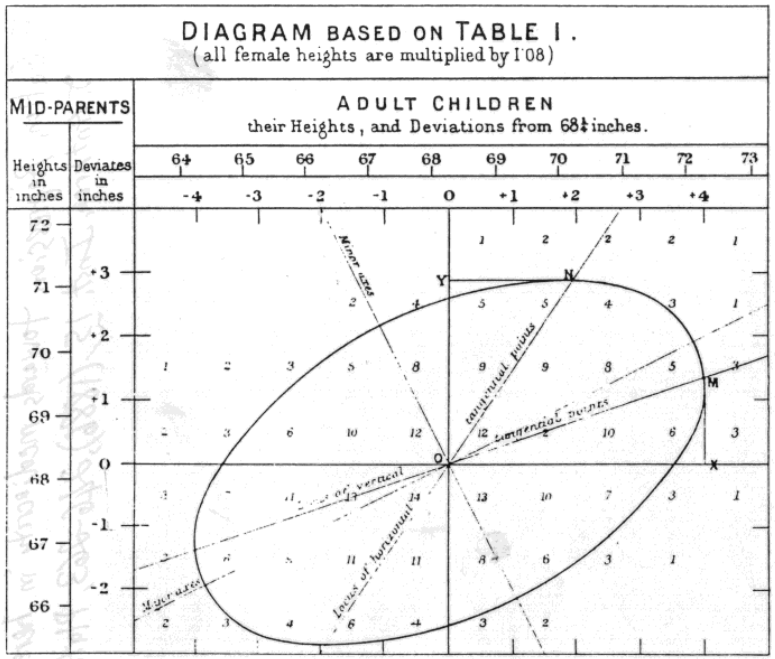
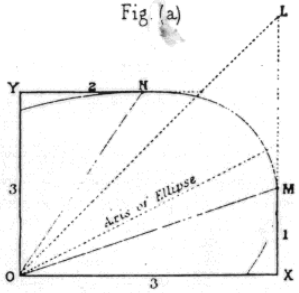


DIAGRAM BASED ON TABLE I.
(all female heights are multiplied by 1·08)
MID-PARENTS
Heights in inches
Deviates in inches
ADULT CHILDREN
their Heights, and Deviations from 68¼ inches.
Minor axes
tangential points
Locus of vertical
Locus of horizontal
Minor axes

Fig. (a)


Axis of Ellipse

at the College of Surgeons, which is, I believe, the tallest skeleton in any museum, that his extraordinary stature of about 7 feet 7 inches would have been a trifle increased if the faces of his dorsal vertebrae had been more parallel and his back consequently straighter.

The beautiful regularity in the statures of a population, whenever they are statistically marshalled in the order of their heights, is due to the due to the number of variable elements of which the stature is the sum. The best illustrations I have seen of this regularity were the curves of male and female statures that I obtained from the careful measurements made at my Anthropometric Laboratory in the International Health Exhibition last year. They were almost perfect.

The multiplicity of elements, some derived from one progenitor, some from another, must be the cause of a fact that has proved very convenient in the course of my inquiry. It is that the stature of the children depends very closely on the average stature of the the two parents, and may be considered in practice as having nothing to do with their individual heights. The fact was proved as follows:– After transmuting the female measurements in the way already explained, I sorted the adult children of those parents who severally differed 1, 2, 3, 4, and 5 or more inches, into separate lines (see Table II). Each line was then divided into similar classes, showing the number of cases in which the children differed 1, 2, 3. &c., inches from the common average of the children in their respective families. I confined my inquiry to large families of six children and upwards, that the common average of each might be a trustworthy point of reference. The entries in each of the different lines were then seen to run in the same way except that in the last of them the children showed a faint tendency to fall into two sets, one taking after the tall parent, the other after the short one; this, however, is not visible in the summary Table II that I annex. Therefore, when dealing with the transmission of stature from parents to children, the average height of the two parents, or, as I prefer to call it, the "mid-parental" height is all we need care to know about them.

TABLE II.

EFFECT UPON ADULT CHILDREN OF DIFFERENCES IN HEIGHT OF THEIR PARENTS.

Difference between the Heights[1] of the Parents in inches.	Proportion per 50 of cases in which the Heights of the Children deviated to various amounts from the Mid-filial Stature of their respective families						Number of Children whose Heights were observed.
	Less than 1 inch.	Less than 1 inch.	Less than 1 inch.	Less than 1 inch.	Less than 1 inch.	Less than 1 inch.	
Under 1	21	35	43	46	48	50	105
1 and under 2	23	37	46	49	50	..	122
2 and under 3	16	34	41	45	49	50	112
3 and under 4	24	35	41	47	49	50	108
5 and above	18	30	40	47	49	50	78

[1] Every female height has been transmuted to its male equivalent by multiplying it by 1.08, and only

those families have been included in which the number of adult children amounted to six, at least.

NOTE.—When these figures are protracted into curves it will be seen—(1) that they run much alike; (2) that their peculiarities are not in sequence; and (3) that the curve corresponding to the first line occupies a medium position. It is therefore certain that differences in the heights of the parents have on the whole an inconsiderable effect on the heights of their offspring.

It must be noted that I use the word parent without specifying the sex. The methods of statistics permit us to employ this abstract term, because the cases of a tall father being married to a short mother are balanced by those of a short father being married to a tall mother. I use the parent to save any complication due to a fact apparently brought out by these inquiries, that the height of the children of both sexes, but especially that of the daughters, takes after the height of the father more than it does after that of the mother. My personal data are insufficient to enable me to speak with any confidence on this point, much less to determine the ratio satisfactorily.

Another great merit of stature as a subject of inquiries into heredity is that marriage election takes little of no account of shortness or tallness. There are undoubtedly sexual preferences for moderate contrast in height, but the marriage choice is guided by so many and more important considerations that questions of stature appear to exert no perceptible influence upon it. This is by no means my only inquiry into this subject, but, as regards the present data, my test lay in dividing the 205 male parents and the 205 female parents into three groups—T, M, and S—that is, tall, medium, and short (medium male measurement being taken as 67 inches and upwards to 70 inches), and in counting the number of marriages in each possible combination between them (see Table III). The result was that men and women of contrasted heights, short and tall or tall and short, married just about as frequently as men and women of similar heights, both tall or both short; there were 32 cases of the one to 27 of the other.

Table III.

S.,t. 12 cases.	M.,t. 20 cases.	T.,t. 18 cases.
S.,m. 25 cases.	M.,m. 51 cases.	T.,m. 28 cases.
S.,s. 9 cases.	M.,s. 28 cases.	T.,s. 14 cases.

Short and tall, $12 + 14 = 32$ cases.

$$\left.\begin{array}{l}\text{Short and short, } 9\\ \text{Tall and tall, } 18\end{array}\right\} = 27 \text{ cases.}$$

In applying the law of probabilities to investigations into heredity of stature, we may therefore regard the married folk as couples picked out of the general population at haphazard.

The advantage of stature as a subject in which the simple laws of heredity may be studied will now be understood. It is a nearly constant value that is frequently measured and recorded, and its discussion is little entangled with considerations of nurture, of

the survival of the fittest, or of marriage selection. We have only to consider the mid-parentage and not to trouble ourselves about the parents separately. The statistical variations of stature are extremely regular, so much that so that their general conformity with the results of calculations based on the the abstract law of frequency of error is an accepted fact by anthropologists. I have made much use of the properties of that law in cross-testing my various conclusions, and always with success. For example, the measure of variability (say the "probable error") of the system of mid-parental heights ought, on the suppositions justified in the preceding paragraphs, to be equal to that of the system of adult male heights, multiplied by the square root of two; this inference is shown to be correct by direct observation.

The only drawback to the use of stature is its small variability. One-half of the population with whom I dealt, varied less than 1.7 inch from the average of all of them, and one-half of the offspring of similar mid-parentages varied less than 1.5 inch from the average of their own heights. On the other hand, the precision of my data is so small, partly due to the uncertainty in many cases whether the height was measured with the shoes on or off, that I find by means of an independent inquiry that each observation, taking one with another, is liable to an error that as often as not exceeds $\frac{2}{3}$ of an inch.

The law that I wish to establish refers primarily to the inheritance of different degrees of tallness and shortness, and only secondarily to that of absolute stature. That is to say, it refers to measurements made from the crown of the head to the level of mediocrity, upwards or downwards as the case maybe, and not from the crown of the head to the ground. In the population with which I deal the level of mediocrity is $68\frac{1}{4}$ inches (without shoes). The same law applying with sufficient closeness both to tallness and shortness, we may include both under the single head of deviations, and I shall call any particular deviation a "deviate." By the use of this word and that of "mid-parentage" we can define the law of regression very briefly. It is that the height-deviate of the offspring is, on the average, two-thirds of the height-deviate of its mid-parentage.

Plate IX, fig. *a*, gives a graphic expression of the data upon which this law is founded. It will there be seen that the relations between the statures of the children and their mid-parents, which are perfectly simple when referred to the scale of deviates at the right hand of the plate, do not admit of being briefly phrased when they are referred to the scale of statures at its left.

If this remarkable law had been based only on experiments on the diameters of the seeds, it might well be distrusted until confirmed by other inquiries. If it were corroborated merely by comparatively small number of observations on human stature, some hesitation might be expected before its truth could be recognised in opposition to the current belief that the child tends to resemble its parents. But more can be urged than this. It is easily to be shown that we we ought to expect filial regression, and that it should amount to some constant fractional part of the value of mid-parental deviation. It is because this explanation confirms the previous observations made both on seeds and on men that I feel justified on the present occasion in drawing attention to this elementary law.

The explanation of it is as follows. The child inherits partly from his parents, partly from his ancestry. Speaking generally, the further his genealogy goes back, the more numerous and varied will his ancestry become, until they cease to differ from any

equally numerous sample taken at haphazard from the race at large. Their mean stature will then be the same as that of the race, in other words, it will be mediocre. Or, to put the same fact into another form, the most probable value of the mid-ancestral deviates in any remote generation is zero.

For the moment let us confine our attention to the remote ancestry and to the mid-parentages, and ignore the intermediate generations. The combination of the zero of the ancestry with the deviate of the mid-parentage is the combination of nothing with something, and the result resembles that of pouring a uniform proportion of pure water into a vessel of wine. It dilutes the wine to a constant fraction of its original alcoholic strength, whatever that strength might have been.

The intermediate generations will each in their degree do the same. The mid-deviate in any one of them will have a value intermediate between that of the mid-parental deviate and the zero value of the ancestry. Its combination with the mid-parental deviate will be as if, not pure water, but a mixture of wine and water in some definite proportion, had been poured into the wine. The process throughout is one of proportionate dilutions, and therefore the joint effect of all of them is to weaken the original wine in a constant ratio. We have no word to express the form of that ideal and composite progenitor, whom the offspring of similar mid-parentages most nearly resemble, and from whose stature their own respective heights diverge evenly, above and below. If he, she, or it, is styled the "generant" of the group, then the law of regression makes it clear that parents are not identical with the generants of their own offspring.

The average regression of the offspring to a constant fraction of their respective mid-parental deviations, which was first observed in the diameters of seeds, and then confirmed by observations on human stature, is now shown to be a perfectly reasonable law which might have been deductively foreseen. It is of so simple a character that I have made an arrangement with pulleys and weights by which the probable average height of the children of known parents can be mechanically reckoned (see Plate IX, fig. *b*). This law tells heavily against the full hereditary transmission of any gift, as only a few of many children would resemble their mid-parentage. The more exceptional the amount of the gift, the more exceptional will be the good fortune of a parent who has a son who equals, and still more if he has a son who overpasses him in that respect. The law is even-handed; it levies the same heavy succession-tax on the transmission of badness as of goodness. If it discourages the extravagant expectations of gifted parents that their children will inherit their powers, it no less discountenances extravagant fears that they will inherit all their weaknesses and diseases.

The converse of this law is very far from being its numerical opposite. Because the most probable deviate of the son is only two-thirds that of his mid-parentage, it does not in the least follow that the most probable deviate of the mid-parentage is $\frac{3}{2}$, or $1\frac{1}{2}$ that of the son. The number of individuals in a population who differ from mediocrity is so preponderant it it is more frequently the case that an exceptional man is the somewhat exceptional son of rather mediocre than the average son of very exceptional parents. It appears from the very same table of observations by which the value of the filial regression was determined when it is read in a different way, namely, in vertical columns instead of in horizontal lines, that the most probable mid-parentage of a man is one that deviates only one-third as much as the man does. There is a great difference between this value of $\frac{1}{3}$ and the numerical converse mentioned above of $\frac{1}{3}$; it is four

and a half times smaller, since $4\frac{1}{2}$, or $\frac{9}{2}$ being multiplied into $\frac{1}{3}$, is equal to $\frac{3}{2}$.

It will be gathered from what has been said, that a mid-parental deviate of one unit implies a mid-grandparental deviate of $\frac{1}{9}$, a mid-ancestral unit in the next generation of $\frac{1}{9}$, and so on. I reckon from these and other data, by methods I cannot stop now to explain, but will do so in the Appendix, that the heritance derived on an average from the mid-parental deviate, independently of what it may imply, or of what may be known concerning the previous ancestry is only $\frac{1}{2}$. Consequently, that similarly derived from a single parent is only $\frac{1}{4}$, and that from a single grandparent is $\frac{1}{16}$.

Let it not be supposed for a moment that any of these statements invalidate the general doctrine that the children of a gifted pair are much more likely to be gifted than the children of a mediocre pair. What they assert is that the ablest child of one gifted pair is not likely to be as gifted as the ablest of all the children of very many mediocre pairs. However, as, notwithstanding this explanation, some suspicion may remain of a paradox lurking in my strongly contrasted results, I will call attention to the form in which the table of data (Table I) was draws up, and give an anecdote connected with it.

It is deduced from a large sheet on which I entered every child's height, opposite to its mid-parental height, and in every case each was entered to the nearest tenth of an in inch. Then I counted the number of entries in each square inch, and copied them out as they appear in the table. The meaning of the table is best understood by examples. Thus, out of a total of 928 children who were born to the 205 mid-parents on my list, there were 18 of the height of 69.2 inches (counting to the nearest inch), who were born to mid-parents of the height of 70.5 inches (also counting to the nearest inch). So again there were 25 children of 70.2 inches born to mid-parents of 69.5 inches. I found it hard at first to catch the full significance of the entries in the table, which had curious relations that were very interesting to investigate. They came out distinctly when I "smoothed" the entries by writing at each intersection of a horizontal column with a vertical one, the sum of the entries in four adjacent squares, and using these to work upon. I then noticed (see Plate X) that lines drawn through entries of the same value formed a series of concentric and similar ellipses. Their common centre lay at the intersection of the vertical and horizontal lines, that corresponded to $68\frac{1}{4}$ inches. Their axes were similarly inclined. The points where each ellipse in succession was touched by a horizontal tangent, lay in a straight line inclined to the vertical in the ratio of $\frac{2}{3}$; those where they were touched by a vertical tangent lay in a straight line inclined to the horizontal in the ration of $\frac{1}{3}$. The same is true in respect of the vertical lines. These and other relations were evidently a subject for mathematical analysis and verification. They were all clearly dependent on three elementary data, supposing the law of frequency of error to be applicable throughout; these data being (1) the measure of racial variability, whence that of mid-parentages may be inferred as has already been explained, (2) that of co-family variability (counting the offspring of like mid-parentages as members of the same co-family), and (3) the average ratio of regression. I noted these values, and phrased the problem in abstract terms such as a competent mathematician could deal with, disentangled from all reference to heredity, and in that shape submitted it to Mr. J. Hamilton Dickson, of St. Peter's College, Cambridge. I asked him kindly to investigate for me the surface of frequency of error that would result from these three data, and the various particulars of its sections, one of which would form the ellipses to which I have alluded.

I may be permitted to say that I never felt such a glow of loyalty and respect towards the sovereignty and magnificent sway of mathematical analysis as when his answer reached me, confirming, by purely mathematical reasoning, my various and laborious statistical conclusions with far more minuteness than I had dared to hope, for the original data ran somewhat roughly, and I had to smooth them with tender caution. His calculation corrected my observed value of mid-parental regression from $\frac{1}{3}$ to $\frac{6}{17.6}$, the relation between the major and minor axis of the ellipses was changed 3 per cent. (it should be as $\sqrt{7} : \sqrt{2}$), their inclination was changed less than 2° (it should be to an angle whose tangent is $\frac{1}{2}$). It is obvious, then, that the law of error holds throughout the investigation with sufficient precision to be of real service, and that the various results of my statistics are not casual and disconnected determinations, but strictly interdependent. In the lecture at the Royal Institution to which I have referred, I pointed out the remarkable way in which one generation was succeeded by another that proved to be its statistical counterpart. I there had to discuss the various agencies of the survival of the fittest, of relative fertility, and so forth; but the selection of human stature as the subject of investigation now enables me get rid of all these complications and to discuss this very curious question under its simplest form. How is it, I ask, that in each successive generation there proves to be the same number of men per thousand, who range between any limits of stature we please to specify, although the tall men are rarely descended from equally tall parents, or the short men from equally short? How is the balance from other sources so nicely made up? The answer is that the process comprises two opposite sets of actions, one concentrative and the other dispersive, and of such a character that they necessarily neutralise one another, and fall into a state of stable equilibrium (see Table IV). By the first set, a system of scattered elements is replaced by another system which is less scattered; by the second set, each of these new elements becomes a centre whence a third system of elements are dispersed.

The details are as follows:—In the first of these two stages we start from the population generally, in the first generation; then the units of the population group themselves, as it were by chance, into married couples, whence the more compact system of mid-parentages is derived, and then by a regression of the values of the mid-parentages the still more compact system of the generants is derived. In the second stage each generant is a centre whence the offspring diverge upwards and downwards to form the second generation. The stability of the balance between the opposed tendencies is due to the regression being proportionate to the deviation. It acts like a spring against a weight; the spring stretches until its resilient force balances the weight, then the two forces of spring and weight are in stable equilibrium; for if the weight be lifted by the hand it will obviously fall down again when the hand is withdrawn, and, if it be depressed by the hand, the resilience of the spring will be thereby increased, so that the weight will rise when the hand is withdrawn.

A simple equation connects the three data of race variability, of the ratio of regression, and of co-family variability, whence, if any two are given, the third may be found. My observations give separate measures of all three and their values fit well into the equation, which is of the simple form—

$$v^2 \frac{p^2}{2} + f^2 = p^2,$$

where $v = \frac{2}{3}$, $p = 1.7$, $f = 1.5$.

It will therefore be understood that the complete table of mid-parental and filial heights may be calculated from two simple numbers, and that the most elementary data upon which it admits of being constructed are—(1) the ratio between the mid-parental and the rest of the ancestral influences, and (2) the measure of the co-family variability.

The mean regression in stature of a population is easily ascertained; I do not see much use in knowing it, but will give the work merely as a simple example. It has already been stated that half the population vary less than 1.7 inch from mediocrity, this being what is technically known as the "probable" deviation. The mean deviation is, by a well-known theory, 1.18 times that of the probable, therefore in this case it is 1.9 inch. The mean loss through regression is $\frac{1}{3}$ of that amount, or a little more than 0.6 inch. That is to say, taking one child with another, the mean amount by which they fall short of their mid-parental peculiarity of stature is rather more than six-tenths of an inch.

The stability of a Type, which I should define as "an ideal form towards which the children of those who deviate from it tend to regress," would I presume, be measured by the strength of its tendency to regress; thus a mean regression from 1 in the mid-parents to $\frac{2}{3}$ in the offspring would indicate only half as much stability as if it had been to $\frac{1}{3}$.

The limits of deviation beyond which there is no regression, but a new condition of equilibrium is entered into, and a new type comes into existence, have still to be explored.

With respect to numerical estimates I wish emphatically to say that I offer them only as being serviceably approximate, though they are mutually consistent, and with the desire that they may be reinvestigated by the help of more abundant and much more accurate measurements than those I have had at command. There are many simple and interesting relations to which I am still unable to assign numerical values for lack of adequate material such as that to which I referred some time back, of the relative influence of the father and the mother on the stature of their sons and daughters.

I do not now pursue the numerous branches that spring from the data I have given, as from a root. I do not speak of the continued domination of one type over others, nor of the persistency of of unimportant characteristics, nor of the inheritance of disease, which is complicated in many cases by the requisite concurrence of two separate heritages, the one of a susceptible constitution, the other of the germs of the disease. Still less do I enter upon the subject of fraternal deviation and collateral descent, which 1 have also worked out.

APPENDIX

I.—*Experiments on Seeds bearing on the Law of Regression*

I sent a set of carefully selected sweet pea seeds to each of several country friends, who kindly undertook to help me. The advantage of sweet peas over other seeds is that they do not cross fertilise, that they are spherical, and that all the seeds in the same pod are of much the same size. They are also hardy and prolific. I selected them as the subject of experiments after consulting eminent botanists. Each packet

TABLE IV.

PROCESS THROUGH WHICH THE DISTRIBUTION OF STATURES, IN SUCCESSIVE GENERATIONS OF THE SAME PEOPLE, REMAINS UNCHANGED.

Height in inches.	Deviation in Inches.	Statistical Distribution of Statures in the several Systems of							
		GENERATION I.		MID-PARENTS.		GENERANTS.		GENERATION II.	
73		2	Concentrates.		Concentrates.		Disperses.	2	
72	+ 4	6		2		1		6	
71	+ 3	10		6		7		10	
70	+ 2	15		16		13		15	Upper Quartile.
69	+ 1	17		26		29		17	
68	0	17		26		29		17	MEDIAN.
67	− 1	15		16		13		15	Lower Quartile.
66	− 2	10	Concentrates.	6	Concentrates.	7	Disperses.	10	
65	− 3	6		2		1		6	
64	− 4	2						2	
Total		100		100		100		100	
Probable derivation ...		1·7		1·2		0·8		1·7	

NOTE.—The cases are symmetrically disposed above and below the common mean value of 68½ inches. The Upper and Lower Quartiles are the values that in each case divide the number of cases above the Median or mean value, and those below it, respectively into equal parts. Thus in each column there are (1) 25 cases per cent. above the Upper Quartile, (2) 25 cases between the Upper Quartile and the Median, (3) 25 cases between the Median and the Lower Quartile, (4) 25 cases below the Lower Quartile. The difference between either Quartile and the Median is technically called the "Probable" deviation.

contained ten seeds of exactly the same weight; those in K being the heaviest, L the next heaviest, and so on down to Q, which was the lightest. The precise weights are given in Table V, together with the corresponding diameter, which I ascertained by laying 100 peas of the same sort in a row. The weights run in an arithmetic series, having a common average difference of 0.172 grain. I do not of course profess to work to thousandths of a grain, though I did to less than tenths of a grain; therefore the third decimal place represents no more thin an arithmetical working value, which has to be regarded in multiplications, lest an error of sensible importance should be introduced by its neglect. Curiously enough, the diameters were found to run approximately in an arithmetic series also, owing, I suppose, to the misshape and corrugations of the smaller seeds, which gave them a larger diameter than if they had been plumped out into spheres. The results are given in Table V, which show that I was justified in sorting the seeds by the convenient method of the balance and weights, and of accepting the weights as directly proportional to the mean diameters, which can hardly be measured satisfactorily except in spherical seeds.

In each experiment seven beds were prepared in partner rows; each was $1\frac{1}{2}$ feet wide and 5 feet long. Ten holes of 1 inch deep were dibbled at equal distances apart along each bed, and one seed was put into each hole. They were then bushed over to keep off the birds. Minute instructions were given and followed to ensure uniformity, which I need not repeat here. The end of all was that the seeds as they became ripe were collected from time to time in bags that I sent, lettered from K to Q, the same letters being stuck at the ends of the beds, and when the crop was coming to an end the whole foliage of each bed was torn up, tied together, labelled, and sent to me. I measured the foliage and the pods, both of which gave results confirmatory of those of the pelts, which will be found in Table VI, the first and last columns of which are those that especially interest us; the remaining columns showing clearly enough how these two were obtained. It will be seen that for each increase of one unit on the part of the parent seed, there is a mean increase of only one-third part of a unit in the filial seed; and again that the mean filial seed resembles the parental when the latter is about 15.5 hundredths of an inch in diameter. Taking then 15.5 as the point towards which filial regression points, whatever may be the parental deviation (within the tabular limits) from that point, the mean filial deviation will be in the same direction, but only one-third as much.

This point of regression is so low that I possessed less evidence than I desired to prove the bettering of the produce of very small seeds. The seeds smaller than Q were such a miserable set that I could hardly deal with them. Moreover, they were very infertile. It did, however, happen that in a few of the sets some of the seeds turned out very well.

If I desired to lay much stress on these experiments, I could make my case considerably stronger by going minutely into the details of the several experiments, foliage and length of pod included, but I do not care to do so.

TABLE V.

WEIGHTS AND DIAMETERS OF SEEDS (SWEET PEA).

Letter of seed.	Weight of one seed in grains.	Length of row of 100 seeds in inches.	Diameter of one seed in hundredths
K	1.750	21.0	21
L	1.578	20.2	20
M	1.406	19.2	19
N	1.234	17.9	18
O	1.062	17.0	17
P	.890	16.1	16
Q	.718	15.2	15

TABLE VI

PARENT SEEDS AND THEIR PRODUCE.

Table showing the proportionate number of seeds (sweet peas) of different seeds produced by parent seeds also of different sizes. The measurements are those of mean diameter, in hundredths of an inch.

Diameter of Parent Seed.	Diameter of filial seeds.								Total.	Mean diameter of Filial Seeds.	
	Under 15	15−	16−	17−	18−	19−	20−	Above 21−		Observed.	Smoothed.
21	22	8	10	18	21	13	6	2	100	17.5	17.3
20	23	10	12	17	20	13	3	2	100	17.3	17.0
19	35	16	12	13	11	10	2	1	100	16.0	16.6
18	34	12	13	17	16	6	2	0	100	16.3	16.3
17	37	16	13	16	13	4	1	0	100	15.6	16.0
16	34	15	18	16	13	3	1	0	100	16.0	15.7
15	46	14	9	11	14	4	2	0	100	15.3	15.4

II.—*Separate Contribution of each Ancestor to the Heritage of the Offspring.*

When we say that the mid-parent contributes two-thirds of his peculiarity of height to the offspring, it is supposed that nothing is known about the previous ancestor. We now see that though nothing is known, something is implied, and that something must be eliminated if we desire to know what the parental bequest, pure and simple, may amount to. Let the deviate of the mid-parent be a, then the implied deviate of the mid-grandparent will be $\frac{1}{3}a$, of the mid-ancestor in the next generation $\frac{1}{9}a$, and so on. Hence the sum of the deviates of all the mid-generations that contribute to the heritage of the offspring is $a(1 + \frac{1}{3} + \frac{1}{9} + \&c.) = a\frac{3}{2}$.

Do they contribute on equal terms, or otherwise? I am not prepared as yet with sufficient data to yield a direct reply, therefore we must try the effects of limiting suppositions. First, suppose they contribute equally; then as an accumulation of ancestral deviates whose sum amounts to $a\frac{3}{2}$, yields an effective heritage of only $a\frac{2}{3}$, it follows that each piece of property, as it were, must be reduced by a succession tax to $\frac{4}{9}$ of its original amount, because $\frac{3}{2} \times \frac{4}{9} = \frac{2}{3}$.

Another supposition is that of successive diminution, the property being taxed afresh in each transmission, so that the effective heritage would be—

$$a\left(\frac{1}{r}+\frac{1}{3r^2}+\frac{1}{3^2r^2}+—\right)=a\left(\frac{3}{3r-1}\right)$$

and this must, as before, be equal to $a\frac{2}{3}$, whence $\frac{1}{r}=\frac{6}{11}$.

The third limiting supposition of a mid-ancestral deviate in any one remote generation contributing more than a mid-parental deviate, is notoriously incorrect. Thus the descendants of "pedigree-wheat" in the (say) twentieth generation show no sign of their mid-ancestral magnitude, but those in the first generation do so most unmistakably. The results of our two valid limiting suppositions are therefore (1) that the mid-parental deviate, pure and simple, influences the offspring to $\frac{4}{9}$ of its amount; (2) that it influences it to the $\frac{6}{11}$ of its amount. These values differ but slightly from $\frac{1}{2}$, and their mean is closely $\frac{1}{2}$, so we may fairly accept that result. Hence the influence, pure and simple, of the mid-parent may be taken as $\frac{1}{2}$, of the mid-grandparent $\frac{1}{4}$, of the mid-great-grandparent $\frac{1}{8}$ and so on. That of the individual parent would therefore be $\frac{1}{4}$, of the individual grandparent $\frac{1}{16}$, of an individual in the next generation $\frac{1}{64}$ and so on.

Explanation of Plates IX and X.

Plate IX, fig. *a*. Rate of Regression in Hereditary Stature.

The short horizontal lines refer to the stature of the mid-parents as given on the scale to the left. These are the same values as those in the left hand column of Table I.

The small circles, one below each of the above, show the mean stature of the children of each of those mid-parents. These are the values in the right hand column of Table I, headed "Medians ." [The Median is the value that half the cases exceed, and the other half fall short of it. It is practically the same as the mean, but is a more convenient value to find, in the way of working adopted throughout in the present instance.]

The sloping line AB passes through all possible mid-parental heights.

The sloping line CD passes through all the corresponding mean heights of their children. It gives the "smoothed" results of the actual observations.

The ratio of CM to AM is as 2 to 3, and this same ratio connects the deviate of every mid-parental value with the mean deviate of its offspring.

The point of convergence is at the level of mediocrity, which is $68\frac{1}{4}$ inches.

The above data are derived from the 928 adult children of 205 mid-parents, female statures having in every case been converted to their male equivalents by multiplying each of them by 1.08.

Fig. *b*. Forecasts of stature. This is a diagram of the mechanism by which the most probable heights of the sons and daughters can be foretold, from the data of the heights of each of their parents.

The weights M and F have to be set opposite to the heights of the mother and father on their respective scales; then the weight sd will show the most probable heights of a son and a daughter on the corresponding scales. In every one of these cases it is the fiducial mark in the middle of each weight by which the reading is to be made. But, in addition to this, the length of the weight sd is so arranged that it is an equal chance (an

even bet) that the height of each son or each daughter will lie within the range defined by the upper and lower edge of the weight, on their respective scales. The length of sd is 3 inches $= 2f$; that is, 2×1.50 inch.

A, B, and C are three thin wheels with grooves round their edges. They are screwed together so as to form a single piece that turns easily on its axis. The weights M and F are attached to either end of a thread that passes over the movable pulley D. The pulley itself hangs from a thread which is wrapped two or three times round the groove of B and is then secured to the wheel. The eight sd hangs from a thread that is wrapped in the same direction two or three times round the groove of A, and is then secured to the wheel. The diameter of A is to that of B as 2 to 3. Lastly, a thread wrapped in the opposite direction round the wheel C, which may have any convenient diameter, is attached to a counterpoise.

It is obvious that raising M will cause F to fall, and *vice versâ*, without affecting the wheels AB, and therefore without affecting sd; that is to say, the parental differences may be varied indefinitely without affecting the stature of the children, so long as their mid-parental height is unchanged. But if the mid-parental height is changed, then that of sd will be changed to $\frac{2}{3}$ of the amount.

The scale of female heights differs from that of the males, each female height being laid down in the position which would be occupied by its male equivalent. Thus 56 is written in the position of 60.48 inches, which is equal to 56×1.08. Similarly, 60 is written in the position of 64.80, which is equal to 60×1.08.

In the actual machine the weights run in grooves. It is also taller and has a longer scale than is shown in the figure, which is somewhat shortened for want of space.

Plate X. This is a diagram based on Table I. The figures in it were first "smoothed" as described in the memoir, then lines were drawn through points corresponding to the same values, just as isobars or isotherms are drawn. These lines, as already stated, formed ellipses. I have also explained how calculation showed that they were true ellipses, and verified the values I had obtained of the relation of their major to their minor axes, of the inclination of these to the coordinates passing through their common centre, and so forth. The ellipse in the figure is one of these. The numerals are not directly derived from the smoothed results just spoken of, but are rough interpolations so as to suit their present positions. It will be noticed that each horizontal line grows to a maximum and then symmetrically diminishes, and that the same is true of each vertical line. It will also be seen that the loci of maxima in these follow the lines ON and OM, which are respectively inclined to their adjacent coordinates at the gradients of 2 to 3, and of 1 to 3. If there had been no regression, but if like bred like, then OM and ON would both have coincided with the diagonal OL, in fig. *a*, as shown by the dotted lines.

I annex a comparison between calculated and observed results. The latter are inclosed in brackets.

Given—

"Probable error" of each system of mid-parentages $= 1.22$.
Ratio of mean filial regression $= \frac{2}{3}$.
"Probable error" of each system of regressed values $= 1.50$.
Sections of surface of frequency parallel to XY are true ellipses.
[Obs.—Apparently true ellipses.]

$MX : YO = 6 : 17.5$, or nearly $1 : 3$.
[Obs.—$1 : 3$.]
Major axes to minor axes $= \sqrt{7} : \sqrt{2} = 10 : 5.35$.
[Obs.—$10 : 5.l$.]
Inclination of major axes to $OX = 26° \, 36'$.
[Obs.—$25°$.]
Section of surface parallel to XY is a true curve of frequency.
[Obs.—Apparently so.]
"Probable error" of that curve $= 1.07$.
[Obs.—1.0 or a little more.]

[*Journal of the Anthropological Institute* **15** (1886), 246–263.]