The Intensity of Natural Selection in Man.

By Karl Pearson, F.R.S.

(Received June 4,—Read June 27, 1912.)

(1) In a paper* communicated to the Royal Society in 1899, and later in greater elaboration published in 'Biometrika,'† 1901, it has been shown on the basis of the inheritance of longevity that the selective death-rate in man amounted to at least 60 per cent. to 80 per cent. of the total death-rate. The matter has been recently reconsidered by Prof. Ploetz,‡ who, dealing with material wholly different from that of Beeton and Pearson came to similar conclusions. The point is a very vital one, for, combined with: (i) the heredity of physical and mental characters in man,§ and (ii) the demonstration that the longer-lived have more offspring,‖ we reach a definite knowledge that Darwinism does apply, and very intensely applies, even to man under civilised conditions.

The difficulty of a direct investigation of the problem lies in securing uniformity of environment. We have to demonstrate that when under the same environment there is a heavier death-rate among a given group of human beings, then among the survivors of this group in a given later period the death-rate will be lessened. Now each group of individuals we attempt to deal with has its own environment, and if that is a bad environment we should expect to find a heavy death-rate both at the earlier and later periods; this obviously must obscure the action of natural selection. For example in districts with a high infant mortality we might expect a high child mortality, say deaths from two to five years of life, because a bad environment sends up the intensity of both. The correlation between deaths in the first year of life (0—1) and in the next four years of life (1—5) for a given district will certainly be positive if no correction be made for varying environment. Quite recently this matter has been discussed by determining the correlation between the ages 0—1 and 1—5 in the administrative counties of England and Wales.¶ As (*a*) the group 0—1 was not followed to 1—5, but the

* 'Roy. Soc. Proc.,' vol. 65, p. 290, *et seq.*

† Vol. 1, p. 50, *et seq.*—especially pp. 74–5.

‡ 'Archiv für Rassen- u. Gesellschafts-Biologie,' 1909, vol. 6, p. 33.

§ "On the Laws of Inheritance in Man. I.—Inheritance of Physical Characters," 'Biometrika,' vol. 2, p. 357; II.—"On the Inheritance of the Mental and Moral Characters in Man," 'Biometrika,' vol. 3, p. 131.

‖ "On the Correlation between Duration of Life and Number of Offspring," 'Roy. Soc. Proc.,' vol. 67, p. 159.

¶ 'Local Government Board Report,' Cd. 5263.

deaths in these age-groups for the *same* years were dealt with, and (*b*) no allowance whatever was made for the differential environment of the administrative counties, it is difficult to find any real bearing of the data on the problem of natural selection in man.

The only method by which data for different districts can be compared is by endeavouring to fix the nature of the environment. We want to know whether under a constant environment the correlation between the death-rates of infancy and of childhood is positive or negative. Dr. E. C. Snow,* working on both English and Prussian data, finds, using a variety of criteria of sameness of environment, that when this factor is allowed for the correlation between infantile and child death-rate is negative, and substantially negative. In other words, within the same group under the same environment, the greater the infantile death-rate, the less is the death-rate among the survivors; that is to say, the physically stronger members of the group are those which survive the ordeal.

Dr. Arthur Newsholme in his recent paper, "National Importance of Child Mortality,"† has asked the question (p. 332):—"Is it certain that a lower infant mortality will produce a survival of an increased proportion of physically inferior children?" and he reaches the conclusion that:—"A high infant death-rate in a given community implies in general a high death-rate in the next four years of life, while low death-rates at both age-periods are similarly associated" (p. 334). "So far from any weeding out of the weaklings being manifest, the counties with a high infant mortality have a death-rate which is relatively higher still in the next four years of life." As evidence for this he cites the correlation between infantile and child death-rates for the same period, and the life-tables for Healthy Districts and for all England and Wales. These life-tables, however, show, as I indicate below, that a heavier infantile mortality has actually a lower child mortality associated with it. It appears to me that the non-allowance for differential environment renders Dr. Newsholme's reasoning insecure.

It is not generally realised that the infantile mortality in England and Wales has not been falling but steadily rising since the restriction in size of families. On the other hand, the child mortality has been steadily falling. The numbers provided by the Registrar-General's Life-Tables‡ are:—

* "The Intensity of Natural Selection in Man," 'Drapers' Company Research Memoirs,' Dulau & Co., 1911.

† 'Journal of the Royal Sanitary Institute,' 1910, vol. 31, p. 334.

‡ 'Supplement to the 65th Annual Report,' 1891—1900, Part I, pp. xlviii—l, London, 1907.

English Life-Tables.

Period.	Males. Deaths per 1000 individuals.		Females. Deaths per 1000 individuals.	
	0—1 years.	1—5 years.	0—1 years.	1—5 years.
1838—1854.....	163·50	134·73	134·71	132·60
1871—1880.....	158·58	127·58	128·73	124·70
1881—1890.....	161·04	104·26	131·13	98·55
1891—1900.....	171·86	94·01	140·66	89·83

Now, these data are the most considerable we can get and they involve all environments in this country. If we might suppose the environment of the country as a whole to have remained constant, we could only conclude that it is certain that a high infant death-rate in a given community implies in general a low death-rate in the next four years of life.

It is difficult to realise how these data have come to be overlooked. It is clear from them that the improved environment of the last 30 to 40 years has not effected any improvement in the infantile death-rate. Either motherhood is less efficient, or the quality of the infants has degenerated, and they can resist worse a better environment. It seems *a priori* unreasonable to suppose the improvement in child mortality wholly due to an improved environment which has produced no effect on infant mortality; it is *a priori* not unreasonable to assume that some of the improved child mortality is due to increased infantile mortality. Can we to any extent determine these proportions?

(2) There are two preliminary points to be considered in this matter. In the first place the actuarial calculation of a life-table is not based on following a given group of persons through life, and determining how many die at each age. At first sight, therefore, our objection (*a*) above applies to approaching the problem of natural selection in man from the life-table standpoint. But a little consideration will show that it is far more justifiable than dealing with local districts in the *same* year. We have in the first place an enormous mass of material—the whole country—and the results are taken for 10 or 16 years. We thus obtain a stable community representing the average of what is taking place in a community gradually changing throughout the epoch. The general experience of actuaries who use such tables for all purposes indicates that the death-rates calculated in this manner for various groups of individuals closely represent what occurs, if the same group be followed through life—the accidental excesses or defects, scarlet fever or measles epidemics, etc., are averaged out when a whole

country and considerable periods are dealt with. I think we may say with fair certitude, for example, that of 1000 male children born in the period 1871–1880, 158·58 would die in the first year of life, and of their survivors, 841 in number, $(841/1000) \times 127{\cdot}58$ would die in the second to the fifth years of life. Further it must be remembered that owing to the period (10 and 16 years) covered, the arrays born in each year but the last one or two have contributed their survivors to die in the whole or part of the child period 1—5.

My next point is the problem of environment. We are told that the environment has been continuously improving during the last 40 to 50 years, and if we ask for a measure of it we are very rightly referred to the falling death-rate, which remained stationary until 1866, and then has been falling in a remarkable way ever since. To correct therefore for continuously improving environment we may take something closely associated with the death-rate; it occurred to me that expectation of life would be an excellent measure of this change of environment. In order, however, to introduce no spurious correlation by taking the expectation of life at birth, which would include the influence of the very mortalities I am dealing with, I have taken the expectation of life at the age of 6 years as the factor by which to correct for the secular change in environment.*

Of course, I very fully realise the audacity of determining correlations from four life-tables only,† but it must be remembered that each one of my figures is based upon a population of millions, and even if we considered our total number of observations four only, the fundamental partial correlations are still immensely significant as compared with their probable errors. Further, I shall show that calculated and observed results are in agreement.

Tables I and II give the data. Underneath the tables I have placed the chief statistical constants; i = infantile death-rate (0—1 years) in deaths per thousand; c = death-rate of children (1—5 years) in deaths per thousand; e = expectation of life in children aged 6.

* Should it be said that the expectation of life at six years of age is influenced by the mortality which occurs in the first five years, this in itself would be to admit that the death-rate is truly selective, the very point we have set out to prove, as against those who hold that the infantile death-rate is not selective.

† We have to bear in mind the vast amount of work involved in computing a table of this kind, and recognise that in calculating 14 life-tables in four series the General Registry Office has achieved a great task.

Table I.—Males.

Years of life-table.	*i.*	*c.*	*e.*
1838—1854	163·50	134·73	49·39
1871—1880	158·58	127·58	50·38
1881—1890	161·04	104·26	52·19
1891—1900	171·86	94·01	52·88
Means	$m_i = 163·744$	$m_c = 115·145$	$m_e = 51·21$
S. D.'s	$\sigma_i = 4·9992$	$\sigma_c = 16·6085$	$\sigma_e = 1·3920$

From these data we further deduce:—

$$r_{ic} = -0·6359$$
$$r_{ie} = +0·5620$$
$$r_{ce} = -0·9941$$

These lead to the partial correlation coefficient of infantile and child death-rates for constant environment

$$_e r_{ci} = -0·8605 \pm 0·0875.$$

Notwithstanding the small number of life-tables dealt with this is most substantially significant. With our measure of constant environment, every increase of the infantile death-rate is accompanied by a marked decrease in the death-rate of the survivors in childhood.

The multiple regression equation is

$$\begin{aligned} c &= 115·145 - 0·3748\,(i - 163·744) - 11·0038\,(e - 51·21) \\ &= 740·021 - 0·3748\,i - 11·0038\,e. \end{aligned} \qquad (1)$$

Thus for constant environment, *i.e.* e constant—

Increase in child death-rate = −0·3748 (increase in infantile death-rate).

Or, looked at in another way—

Percentage decrease in child death-rate = 0·533 (percentage increase in infantile death-rate).

In other words if you increase the infantile death-rate by 10 per cent. you decrease the child death-rate by 5·3 per cent.

The approximation of the method is indicated by the following table:—

Period.	Observed *c.*	Calculated *c.*	Δ.
1838—1854	134·73	135·26	+0·53
1871—1880	127·58	126·21	−1·37
1881—1890	104·26	105·38	+1·12
1891—1900	94·01	93·73	−0·28

The results are in far better accord than we might have anticipated, and we see from these male data that—

(i) Our environmental measure appears to be justified by its close correlation (0·994) with the child death-rate.

(ii) Notwithstanding this the infantile death-rate is fairly closely correlated with the environmental measure in the *negative* sense.

(iii) The infantile and child death-rates, correcting for the environmental factor, are very substantially negatively correlated.

We may now turn to the female data—

Table II.—Females.

Years of life-table.	i.	c.	e.
1838—1854	134·71	132·60	50·00
1871—1880	128·73	124·70	52·56
1881—1890	131·13	98·55	54·35
1891—1900	140·66	89·83	55·18
Means	$m_i = 133·81$	$m_c = 111·42$	$m_e = 53·02$
S. D.'s	$\sigma_i = 4·4922$	$\sigma_c = 17·7251$	$\sigma_e = 1·9853$

From these data we further deduce

$$r_{ic} = -0·5080$$
$$r_{ie} = +0·3042$$
$$r_{ce} = -0·9507$$

The partial correlation coefficient between infantile and child death-rates is

$$_e r_{ic} = -0·7399 \pm 0·1459.$$

This is again substantial and significant. It means again that by fixing the environment as far as lies in our power, we have increased the negative correlation between infantile and child death-rates.

The multiple regression formula is now

$$\begin{aligned} c &= 111·42 - 0·9514\,(i - 133·81) - 7·8335\,(e - 53·02) \\ &= 654·06 - 0·9514\,i - 7·8335\,e. \end{aligned} \qquad (2)$$

Clearly for a constant environment—

Increase in child death-rate = −0·9514 (increase in infantile death-rate).

Or, in other words—

Percentage decrease in child death-rate = 1·1424 (increase in infantile death-rate).

Thus in the case of females the child death-rate goes down about 1 per cent. for every rise of 1 per cent. in the infantile death-rate.

As we have seen, the correlations are not as high in the case of the females as in that of the males. This is probably, to some extent, due (i) to the ages of the women being less reliable than the ages of the men, and this would affect e, and (ii) to the lower infantile death-rate in women, the selection being very likely pushed to a higher age; the mere difficulties of birth are greater in the case of the boy and selection may thus be more immediate and stringent in his case. The lower values of the correlations make the agreements between observed and calculated values less close, still there is nothing much to complain of here:—

Period.	Observed.	Calculated.	Δ.
1838—1854	132·60	134·22	+1·62
1871—1880	124·70	119·86	−4·84
1881—1890	98·55	103·55	+5·00
1891—1900	89·93	87·98	−1·85

The remarkable feature of these English life-tables has been the falling child death-rate accompanying the rising infantile death-rate, a phenomenon which should have led those who assert that a high infantile death-rate implies a high death-rate in the next four years of life to pause. It is noteworthy that the Registrar-General has drawn up life-tables in six cases for "selected material," *i.e.* for "selected healthy districts."* These tables give the following results:—

Period.	Males.			Females.		
	i.	*c.*	*e.*	*i.*	*c.*	*e.*
1849—1853 ...	112·80	82·40	53·94	92·64	79·13	53·48
1881—1890 ...	119·15	61·30	56·37	93·50	56·62	57·31
1891—1900 ...	121·50	51·31	57·52	95·08	48·32	58·80

Without venturing to work out anything on the basis of two sets of three tables, we can yet recognise from the crude numbers that precisely the same phenomena occur even in the "healthy districts," *i.e.* improving environment is related to an increasing not a decreasing infantile mortality and while the child mortality decreases with better conditions, it is highly correlated in a negative sense with infantile mortality, a correlation which will not be

* 'Supplement to the 65th Annual Report,' Part I, pp. lviii—cxi, 1907.

reduced but emphasized when correction is made for the secular change in environment.*

It is not my purpose to enter into a discussion of the rise in infantile mortality which has gone on in this country since the fall in the birth-rate started, but I think enough has been done to indicate that when allowance is made for environment, a heavy infantile death-rate indicates a reduced child death-rate. Further, so far from the conclusion that "a high infant death-rate in a given community implies in general a high death-rate in the next four years of life" being true, it is quite incorrect for 14 of the most important English life-tables. These tables support parallel evidence of other kinds, that the Darwinian theory has application to civilised man, and that a heavy death-rate does mean the elimination of the weaklings. To recognise this as a scientific law which controls the evolution of man as of other species is one thing; to assert that the men of science who accept it desire themselves or through the State to play Herod to our modern infants is another thing and a pernicious thing. It is the duty of science to discover what is happening in the first place, and there appears to me a fair amount of evidence now to show that Darwinism does apply to man, and that, for a constant environment, the higher the infantile death-rate, the more resistant will be the surviving child population.

* Because the correlation between c and e is so nearly perfect that the term $\sqrt{(1-r_{ce}^2)}$ in the denominator of ${}_e r_{ci}$ dominates the relationship.