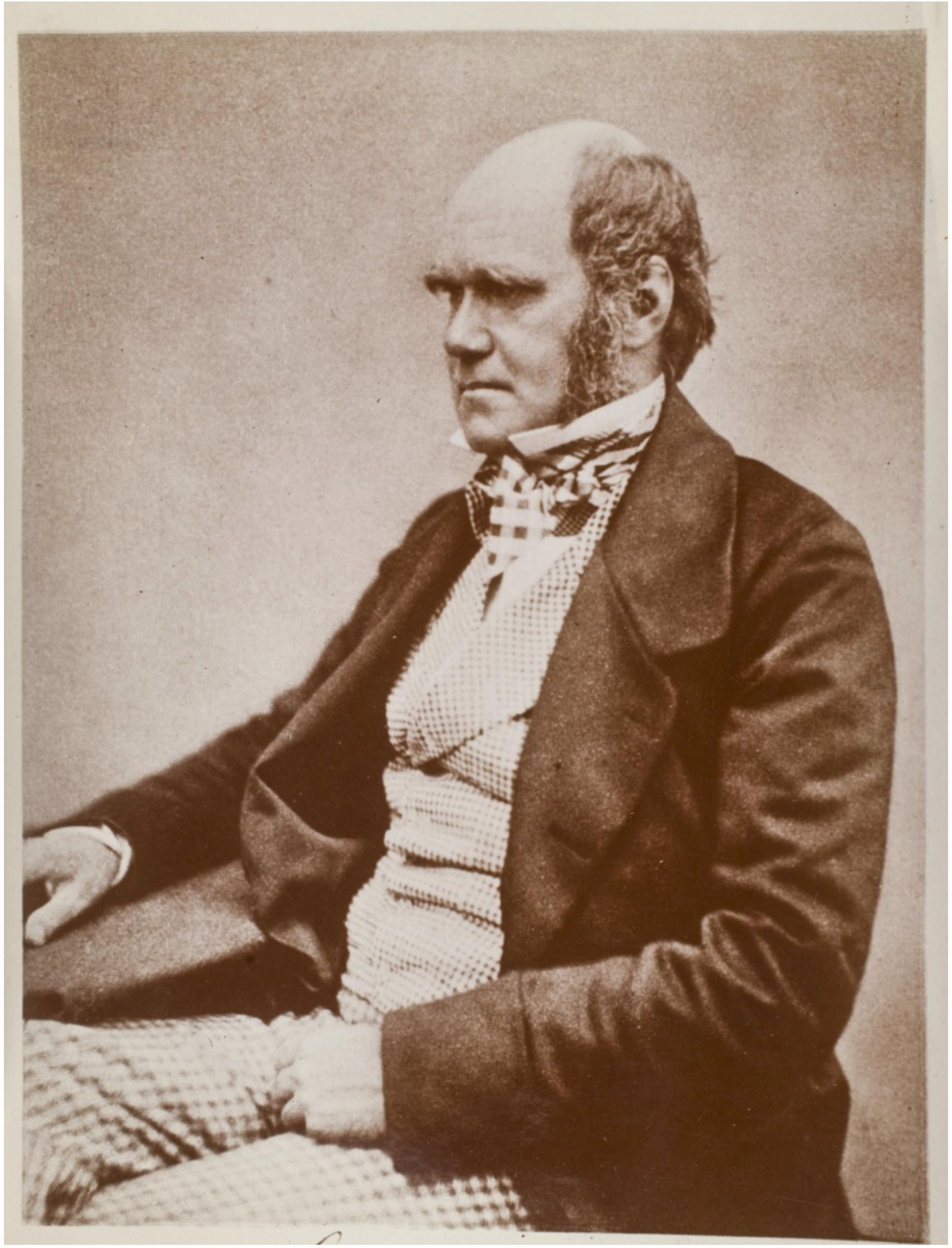
The Causes of Variation

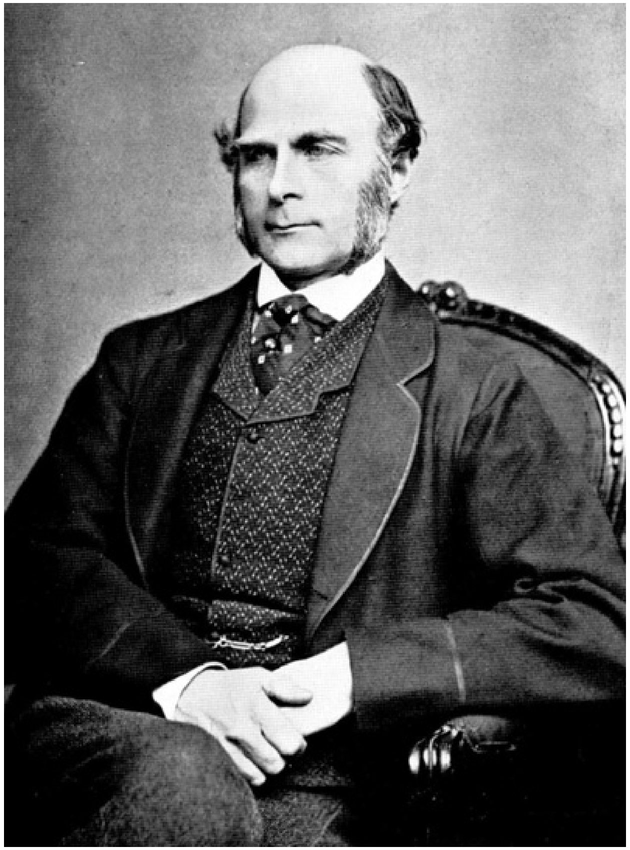
Twin Methodology Workshop
Boulder,
March 2010

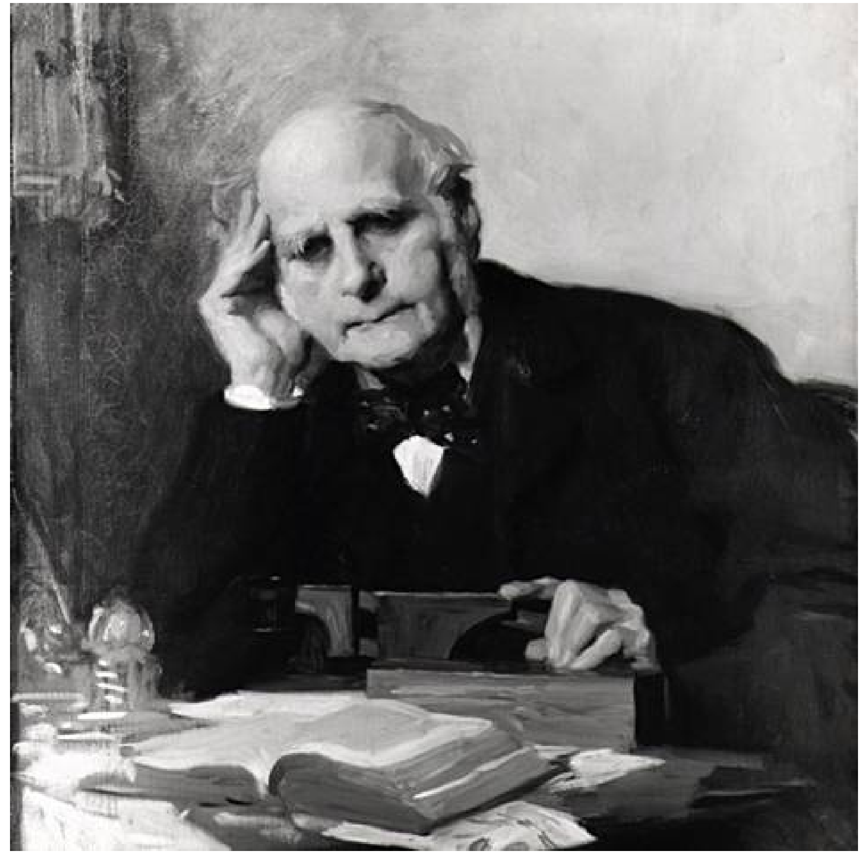
Don’t Panic!

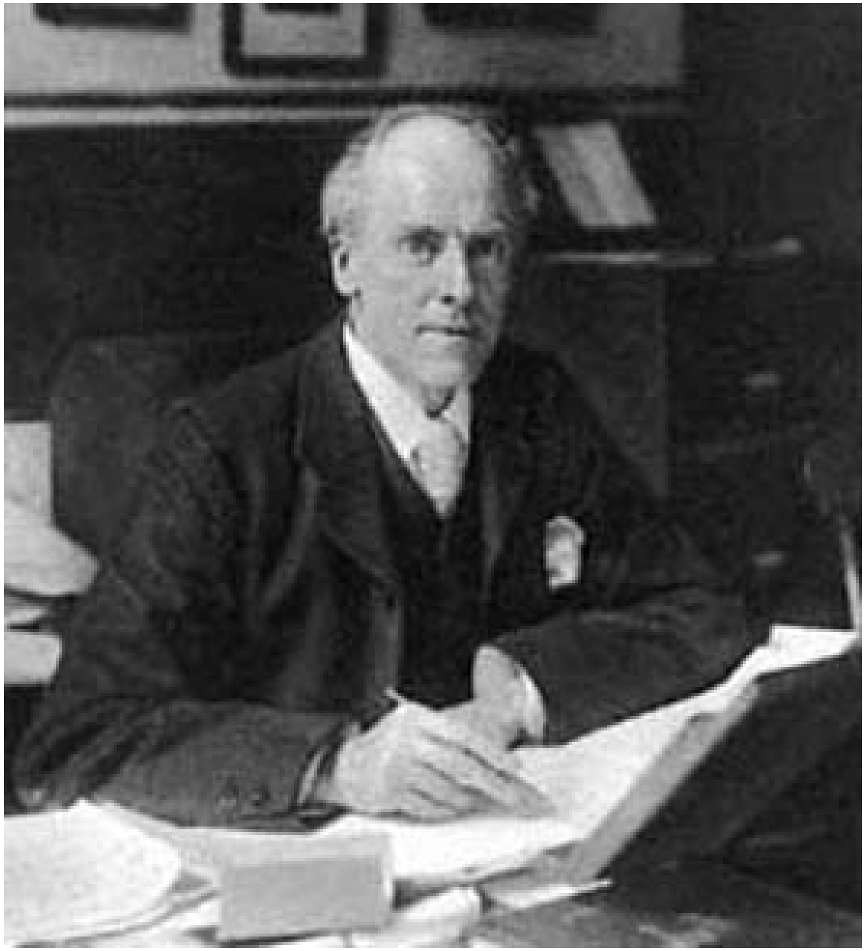
Pretest

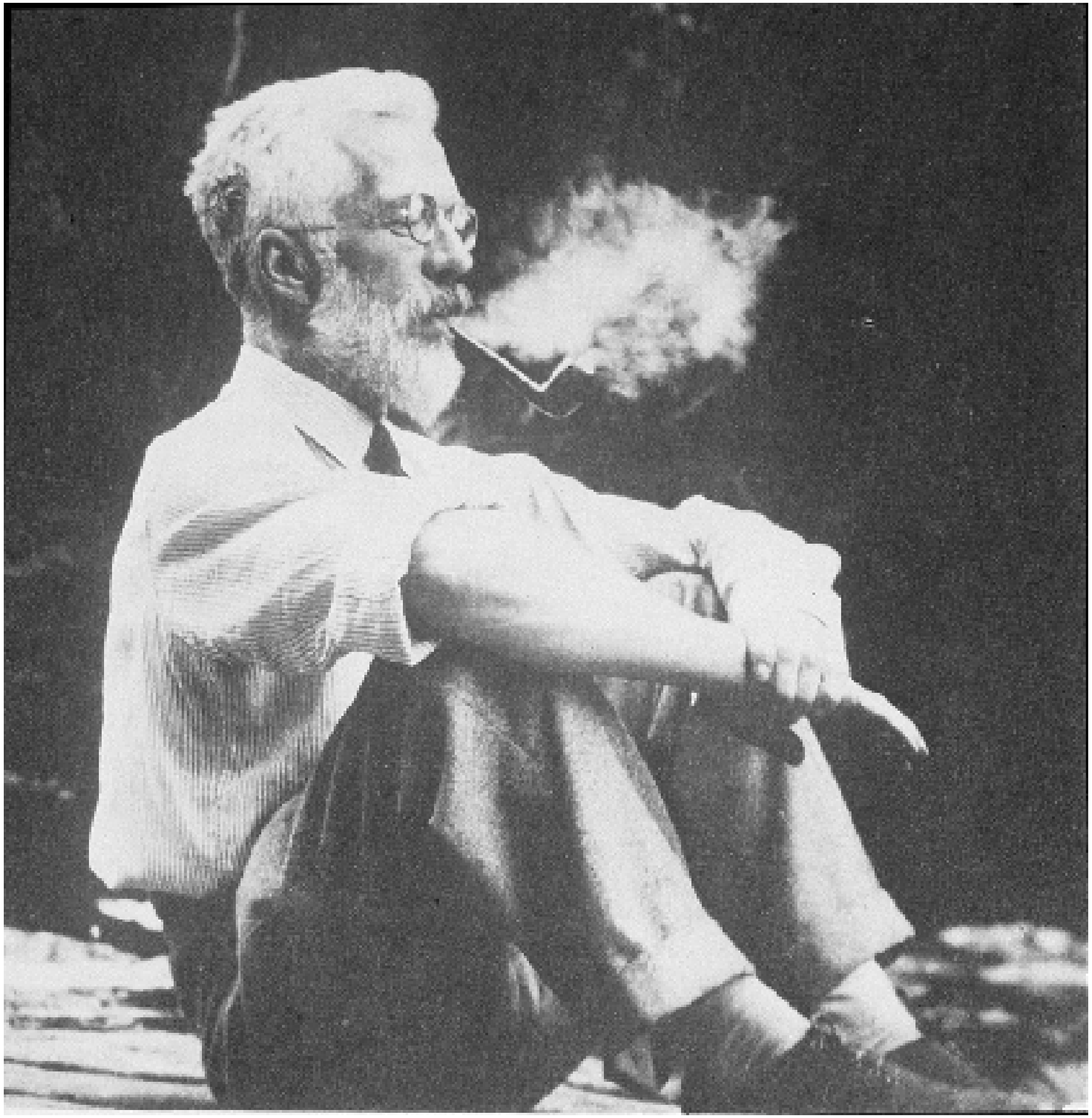
Name the Following People

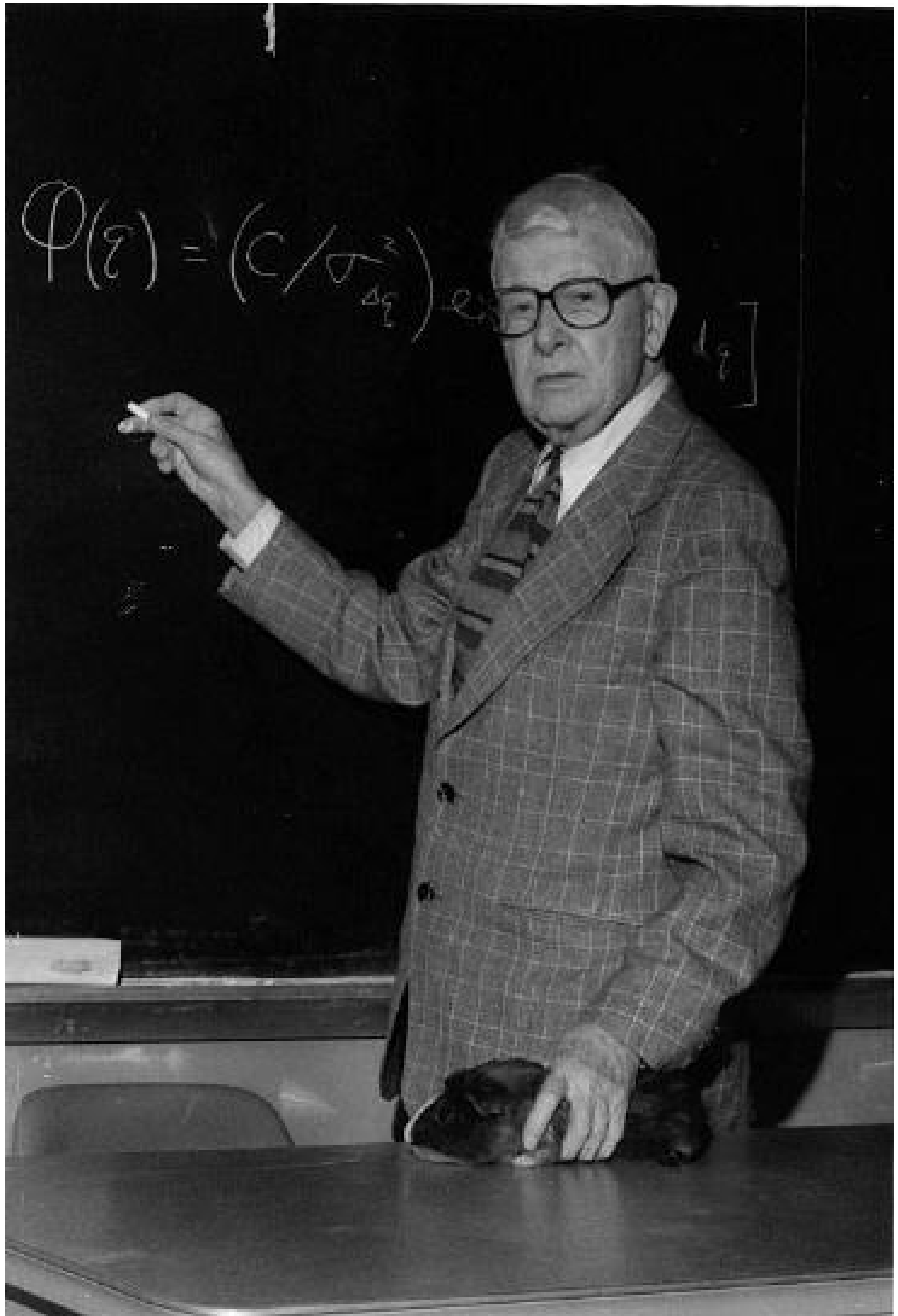













Extra Credit:

1) What is the guy in the middle pointing to?
2) What is he saying?

“Genetics”

The Study of

Variation and Heredity

“Variation”

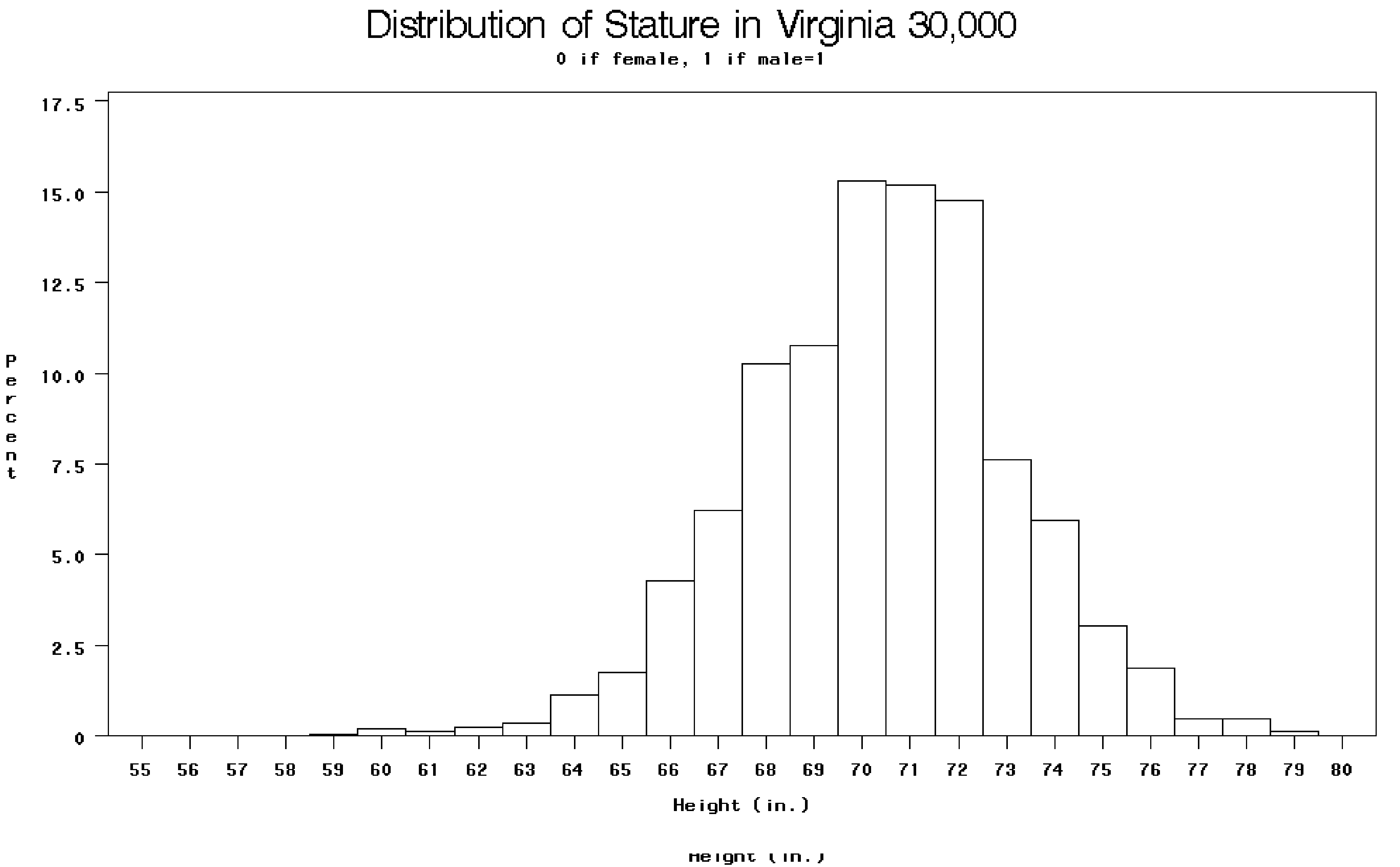
“Why aren’t we all the same?”

“Heredity”

“Why do things run in families?”

“VARIATION”

Continuous variation


Distribution of Stature in Virginia 30,000
0 if female, 1 if male=1
Percent
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
Height (in.)
Height (in.)

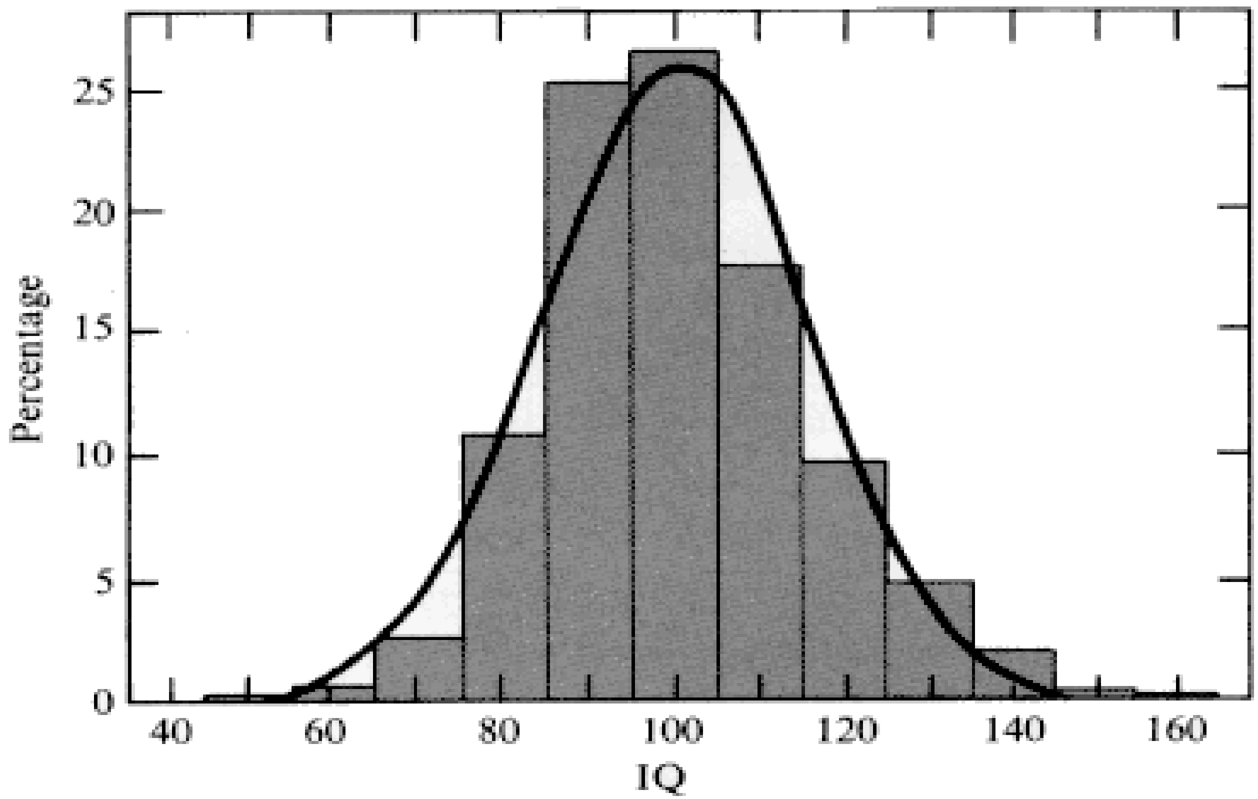
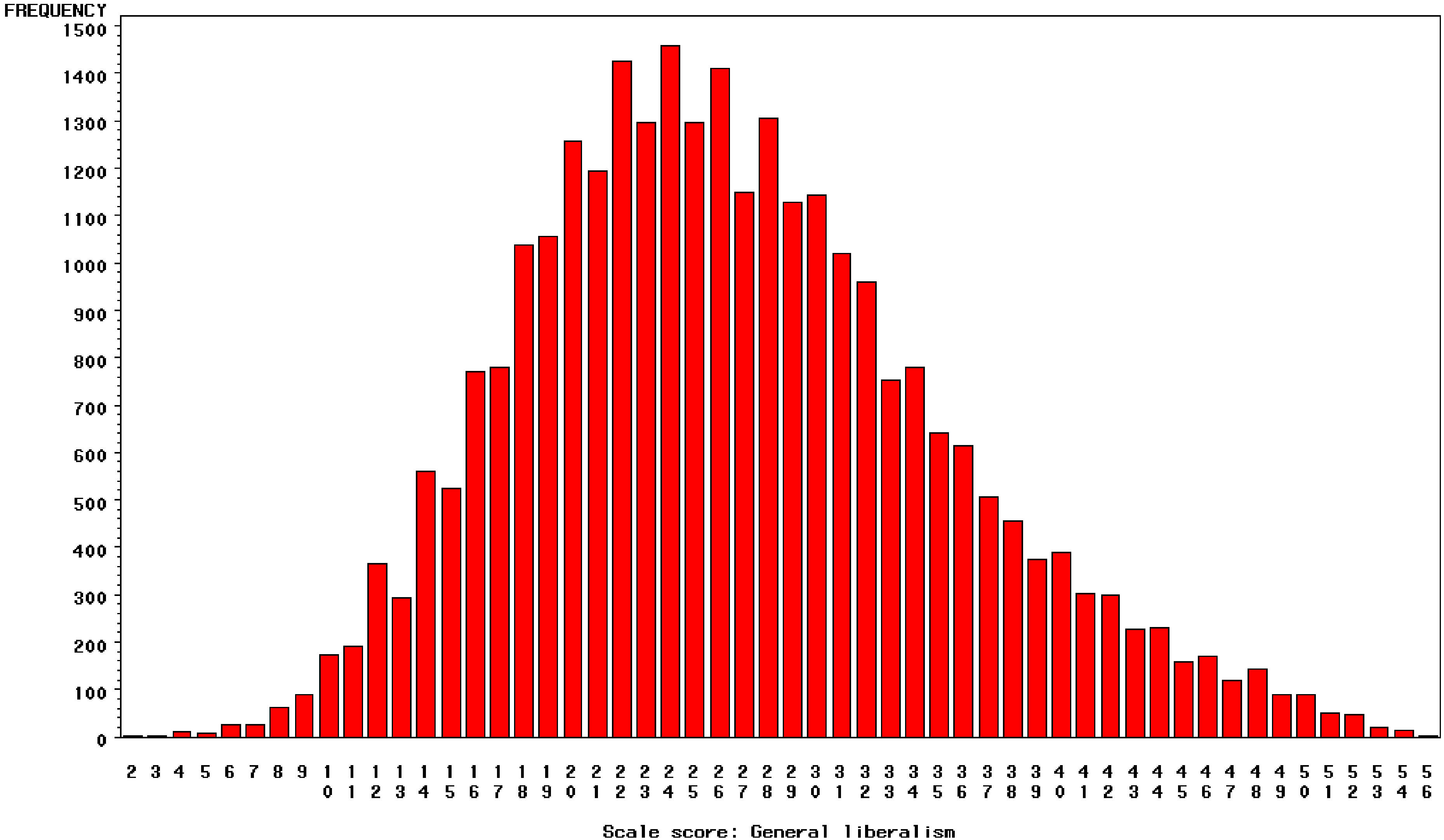


FIGURE 9.4
The distribution of IQ among the 14,963 children born in Scotland on February 1, May 1, August 1, and November 1, 1926. The shaded histogram shows the percentages of the group with IQ's in various ranges of 10 points. This grouping is artificial and is done solely for ease of representation: it does not imply any discontinuity in the values of IQ that children can show. The continuous curve shows the ideal distribution calculated from the observations and representing the statistical population of which the children actually observed are regarded as forming a sample. (Data from MacMeekan; from Mather 1964.)

“Liberalism”



FREQUENCY
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 56
Scale score: General liberalism

Categorical Outcomes

Often called “threshold traits” because people “affected” if they fall above some level (“threshold”) of a measured or hypothesized continuous trait.

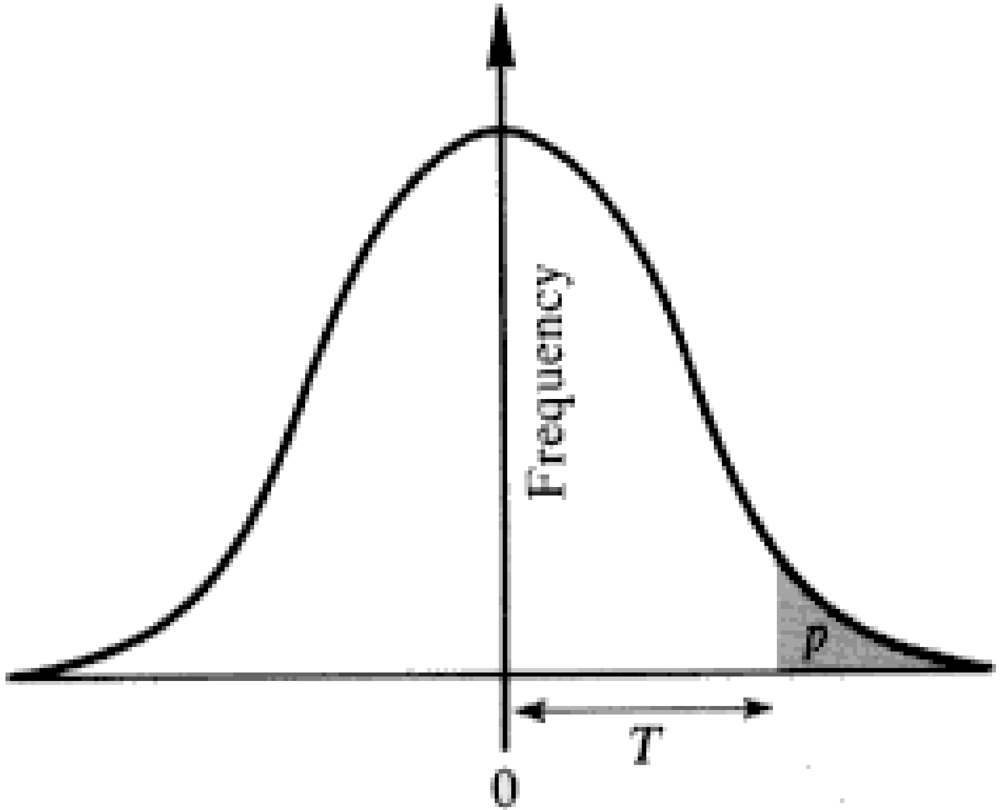
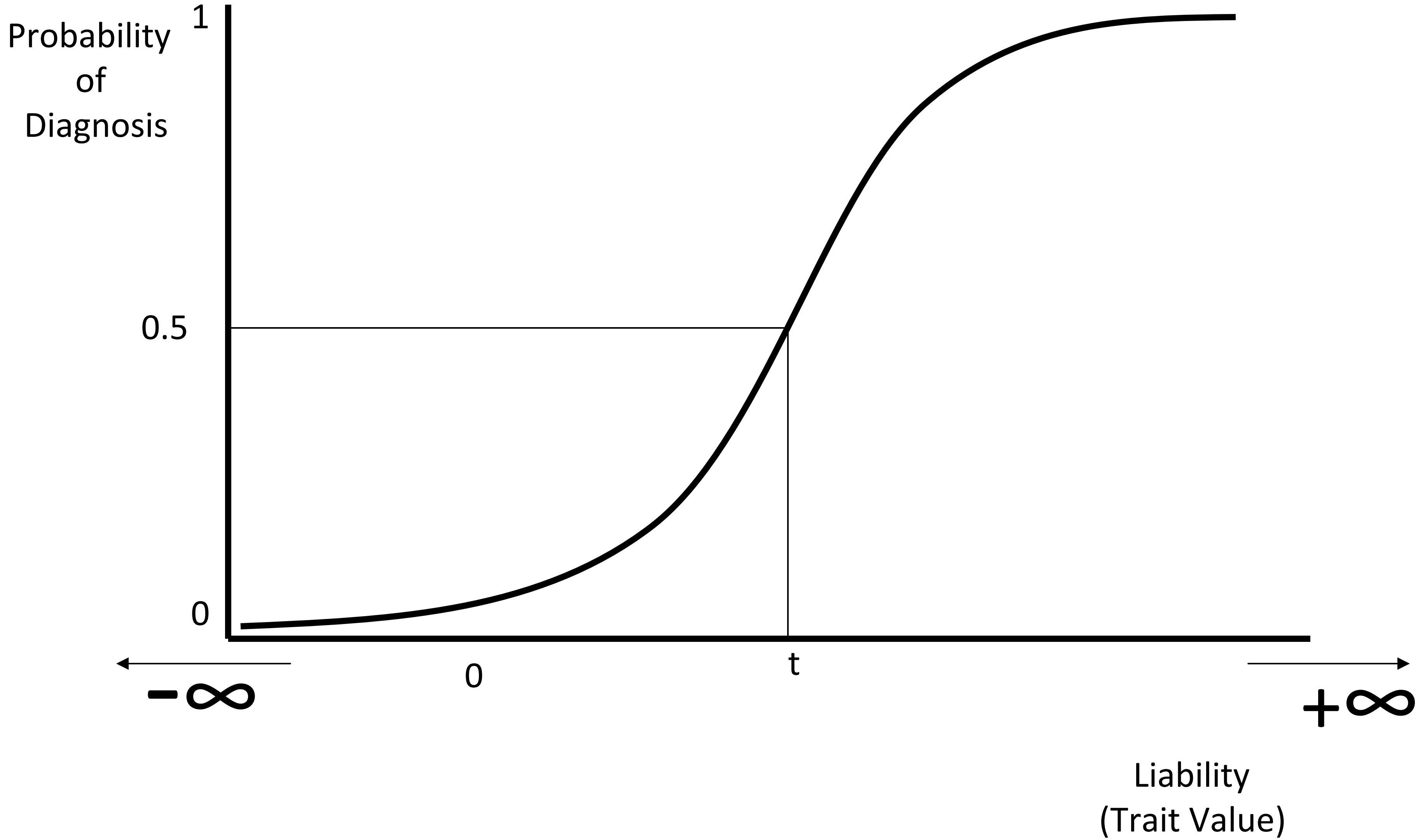


FIGURE 9.5
Threshold model. All individuals with a value of x greater than T are affected. The proportion of affected individuals is the area under the distribution curve beyond T.

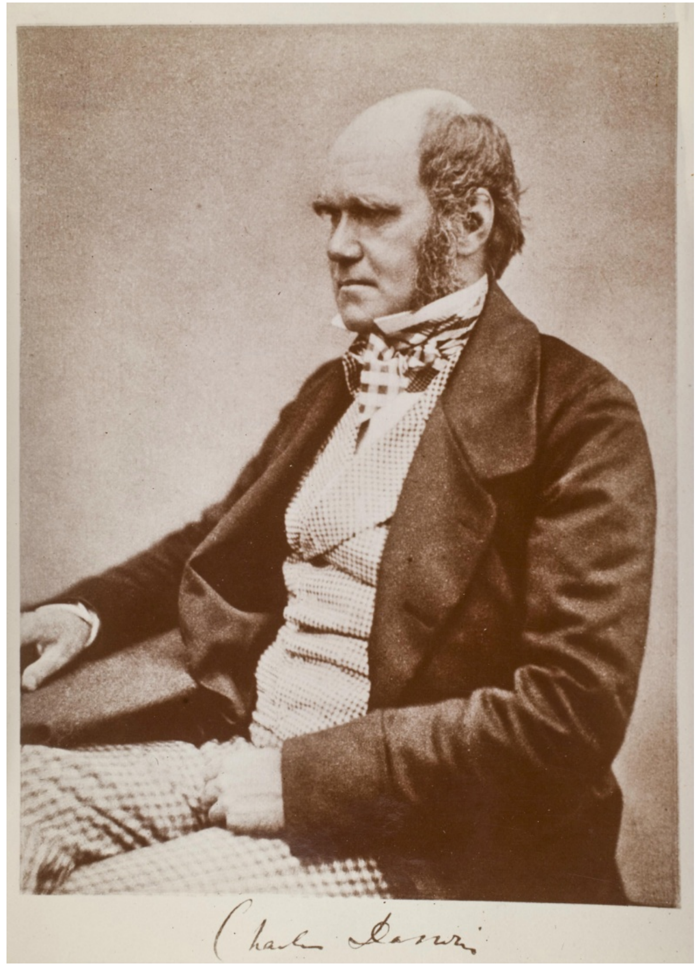
∞

Relationship between continuous normal "liability" and risk of "diagnosis"



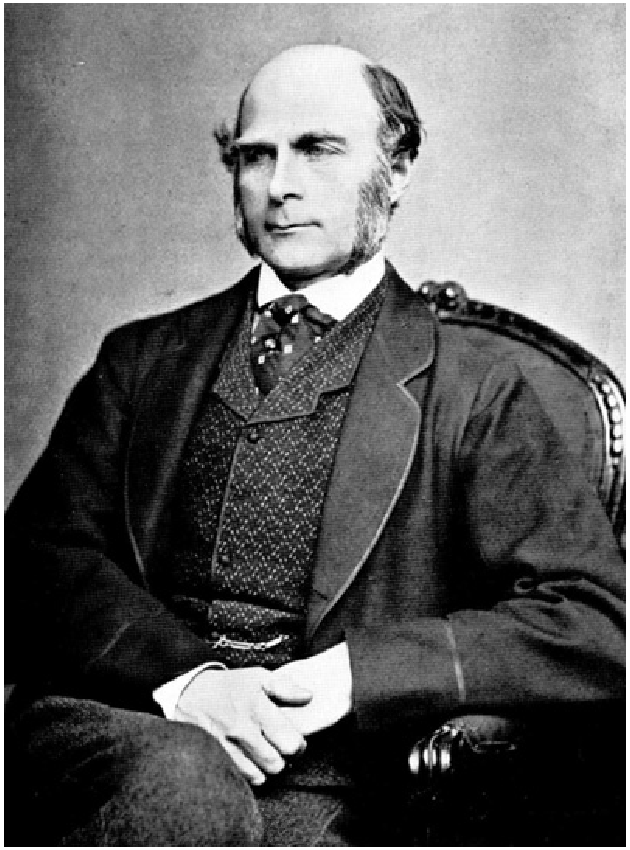
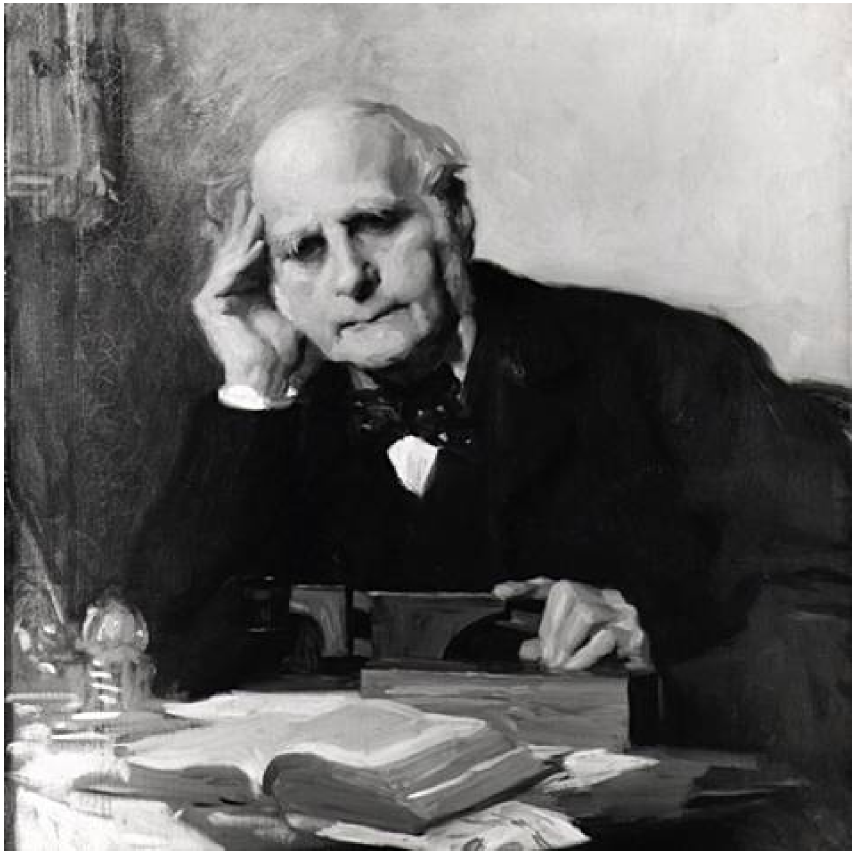
“HEREDITY”

Charles Darwin (1809-1882)



1865: On the Origin of Species

Francis Galton (1822-1911)

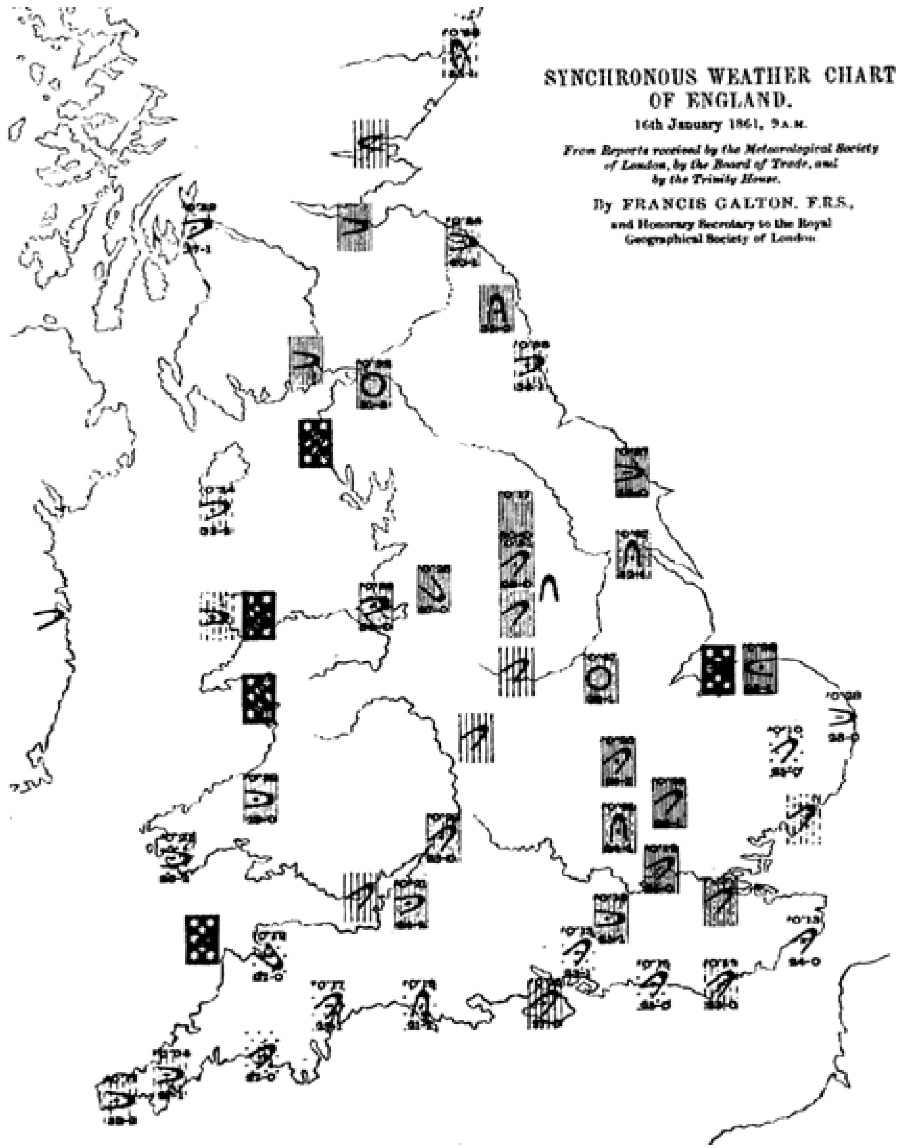
 

1869: Hereditary Genius

1883: Inquiries into Human Faculty and its Development

1884-5: Anthropometic Laboratory at “National Health Exhibition”

Galton’s Other Work e.g. Meteorology



Hereditary Genius (1869, p 317)

	Judges, p. 61.	Statesmen, p. 109.	Commanders, p. 148.	Literary, p. 171.	Scientific, p. 195.	Poets, p. 227.	Artists, pp. 238 and 249.	Divines, p. 275.	Illustrious and Eminent Men of all Classes.		
	B.	B.	B.	B.	B.	B.	B.	B.	B.	C.	D.
Father	26	33	47	48	26	20	32	28	31	100	31
Brother	35	39	50	42	47	40	50	36	41	150	27
Son	36	49	31	51	60	45	89	40	48	100	48
Grandfather	15	28	16	24	14	5	7	20	17	200	8
Uncle	18	18	8	24	16	5	14	40	18	400	5
Nephew	19	18	35	24	23	50	18	4	22	400	5
Grandson	19	10	12	9	14	5	18	16	14	200	7
Great-grandfather	2	8	8	3	0	0	0	4	3	400	1
Great-uncle	4	5	8	6	5	5	7	4	5	800	1
First cousin	11	21	20	18	16	0	1	8	13	800	2
Great-nephew	17	5	8	6	16	10	0	0	10	800	1
Great-grandson	6	0	0	3	7	0	0	0	3	400	1
All more remote	14	37	44	15	23	5	18	16	31	?	...

Galton's Anthropometric Laboratory:

ANTHROPOMETRIC LABORATORY

For the measurement in various ways of Human Form and Faculty.

Entered from the Science Collection of the S. Kensington Museum.

This laboratory is established by Mr. Francis Galton for the following purposes:—

1. For the use of those who desire to be accurately measured in many ways, either to obtain timely warning of remediable faults in development, or to learn their powers.

2. For keeping a methodical register of the principal measurements of each person, of which he may at any future time obtain a copy under reasonable restrictions. His initials and date of birth will be entered in the register, but not his name. The names are indexed in a separate book.

3. For supplying information on the methods, practice, and uses of human measurement.

4. For anthropometric experiment and research, and for obtaining data for statistical discussion.

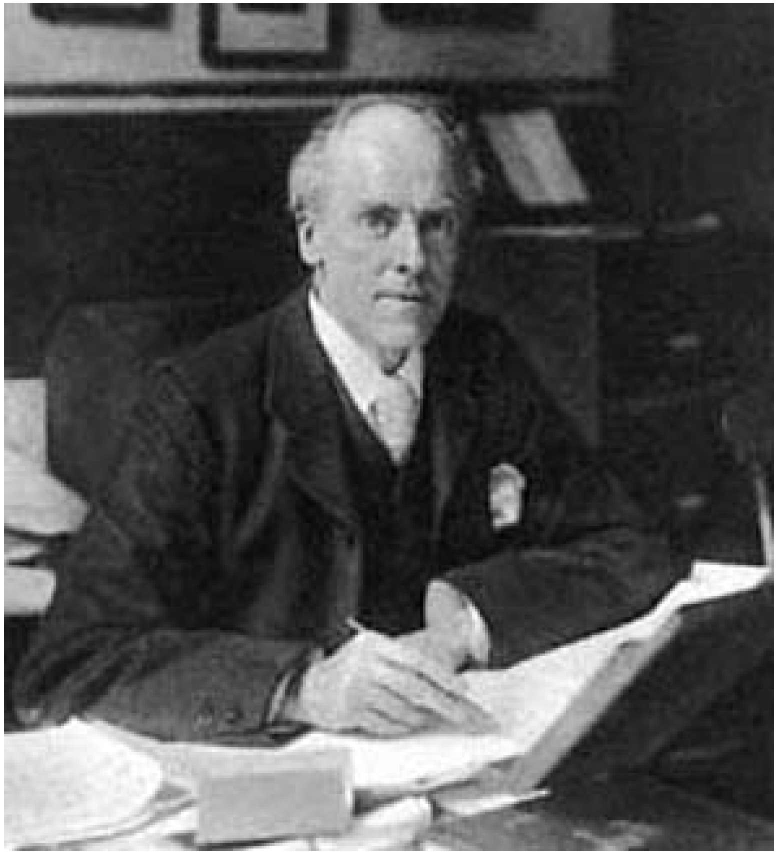
Charges for making the principal measurements:
THREEPENCE each, to those who are already on the Register.
FOURPENCE each, to those who are not:— one page of the Register will thenceforward be assigned to them, and a few extra measurements will be made, chiefly for future identification.

The Superintendent is charged with the control of the laboratory and with determining in each case, which, if any, of the extra measurements may be made, and under what conditions.

H & W. Brown, Printers, 20 Fulham Road, S.W.

Francis Galton's First Anthropometric Laboratory at the International Health Exhibition, South Kensington, 1884–5.

Karl Pearson (1857-1936)



1903: On the Laws of Inheritance in Man: I Physical Characteristics (with Alice Lee)

1904: II Mental and Moral Characteristics

1914: The Life, Letters and Labours of Francis Galton

FAMILY MEASUREMENTS.

Professor KARL PEARSON, of University College, London, would esteem it a great favour if any persons in a position to do so, would assist him by making one set (or if possible several sets) of anthropometric measurements on their own family, or on families with whom they are acquainted. The measurements are to be made use of for testing theories of heredity, no names, except that of the recorder, are required, but the Professor trusts to the *bona fides* of each recorder to send only correct results.

Each family should consist of a father, mother, and at least one son or daughter, not necessarily the eldest. The sons or daughters are to be at least 18 years of age, and measurements are to be made on not more than two sons and two daughters of the same family. If more than two sons or two daughters are easily accessible, then not the tallest but the eldest of those accessible should be selected.

To be of real service the whole series ought to contain 1000—2000 families, and therefore the Professor will be only too grateful if anyone will undertake several families for him.

Copies of this paper, together with cards for recording data, may be obtained from

....

or from the above-named Professor.

Pearson and Lee’s diagram for measurement of “span” (finger-tip to finger-tip distance)

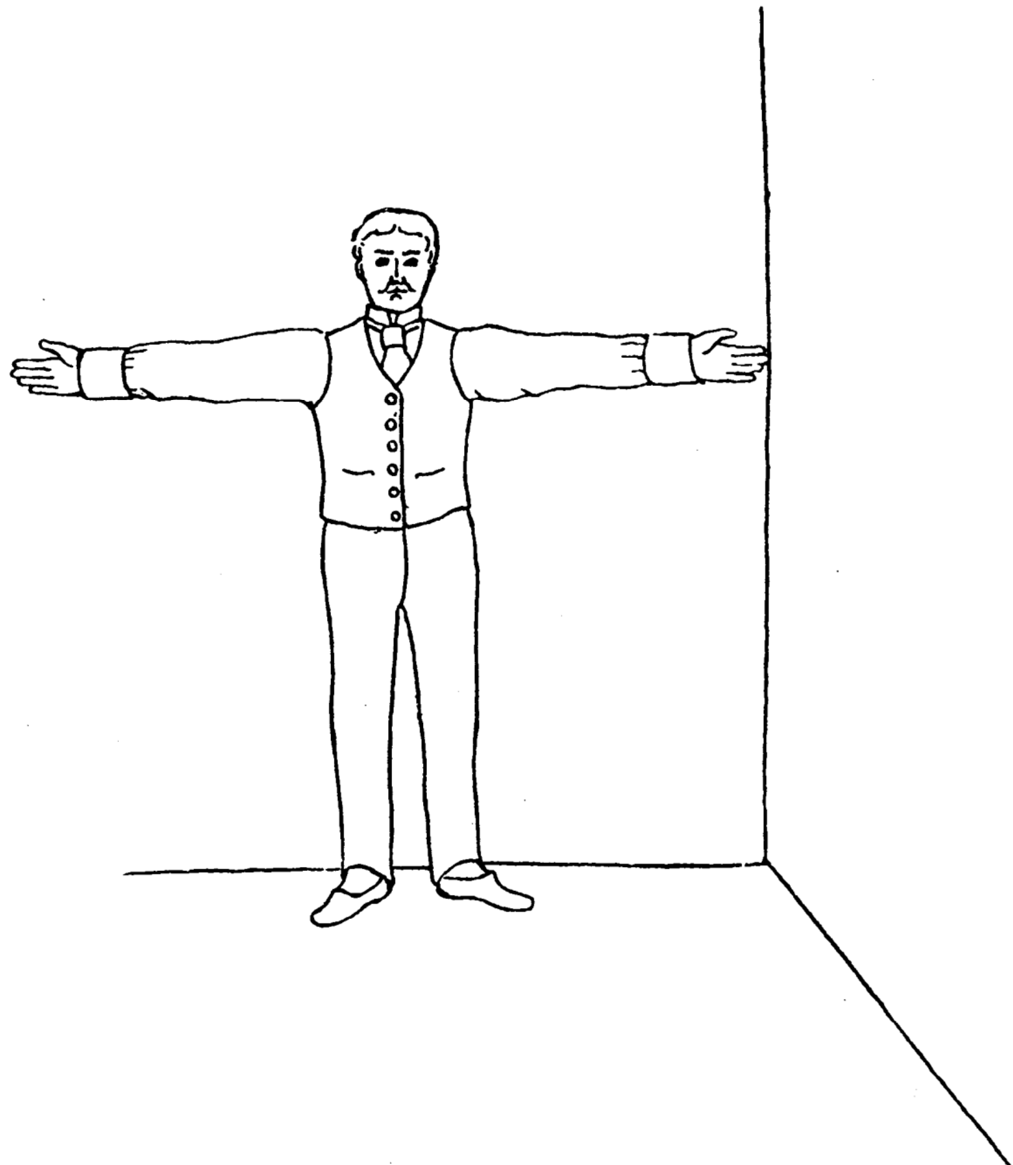
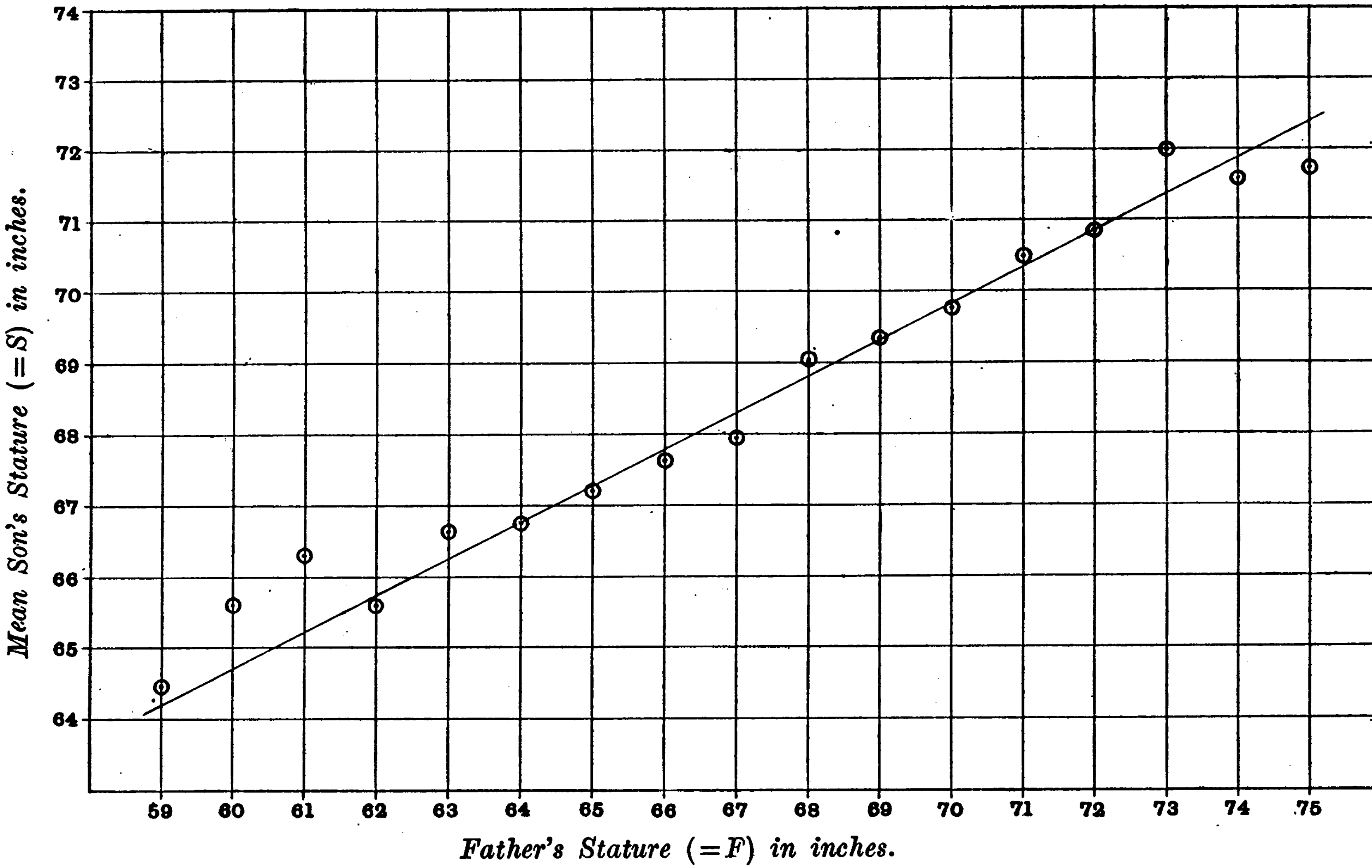


DIAGRAM I. *Probable Stature of Son for given Father's Stature.*

Regression Line: $S = 33{\cdot}73 + {\cdot}516\,F$. 1078 Cases.



From Pearson and Lee (1903) p.378

TABLE IV.

Coefficients of Heredity. Parents and Offspring.

Character	Father and		Mother and	
	Son	Daughter	Son	Daughter
Stature	·514 ± ·015	·510 ± ·013	·494 ± ·016	·507 ± ·014
Span	·454 ± ·016	·454 ± ·014	·457 ± ·016	·452 ± ·015
Forearm	·421 ± ·017	·422 ± ·015	·406 ± ·017	·421 ± ·015

From Pearson and Lee (1903) p.378

Correlation Coefficients for Direct Fraternal Heredity.

Character	Brother and Brother	Sister and Sister	Brother and Sister	Mean
Stature Span Forearm	·511 ± ·028 ·549 ± ·026 ·491 ± ·029	·537 ± ·022 ·555 ± ·021 ·507 ± ·023	·553 ± ·013 ·525 ± ·013 ·440 ± ·015	·534 ·543 ·479
Mean	·517	·533	·506	·519
Eye Colour*	·517 ± ·020	·446 ± ·023	·462 ± ·022	·475
Total mean	·517	·511	·495	·508

From Pearson and Lee (1903) p.387

Assortative Mating. Based on 1000 to 1050 Cases of Husband and Wife.

	Husband's Character	Wife's Character	Correlation and Probable Error	Symbol
Direct	Stature	Stature	·2804 ± ·0189	r_{12}
	Span	Span	·1989 ± ·0204	r_{34}
	Forearm	Forearm	·1977 ± ·0205	r_{56}

From Pearson and Lee (1903) p. 373

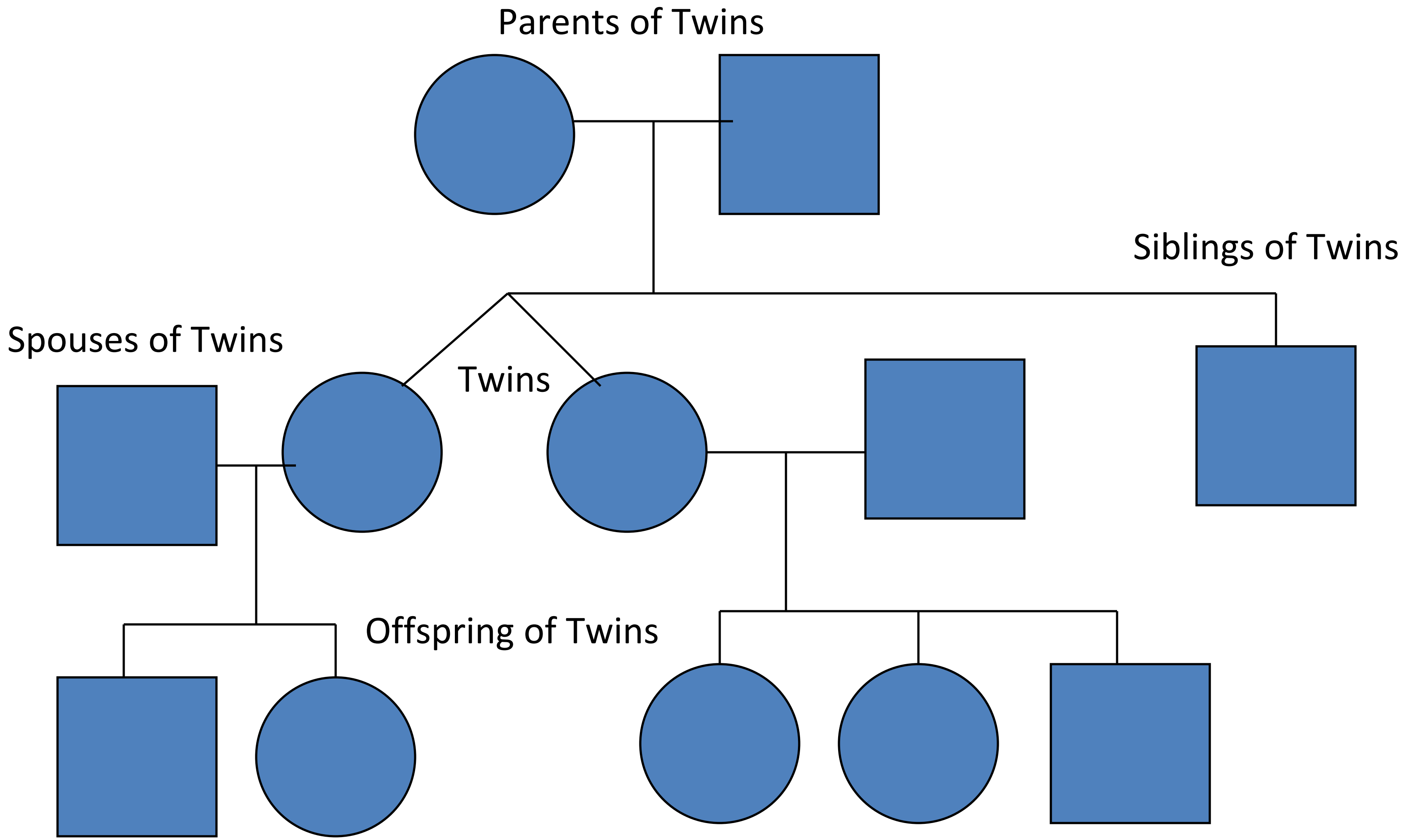
Modern Data

The Virginia 30,000
(N=29691)

The Australia 22,000

(N=20480)

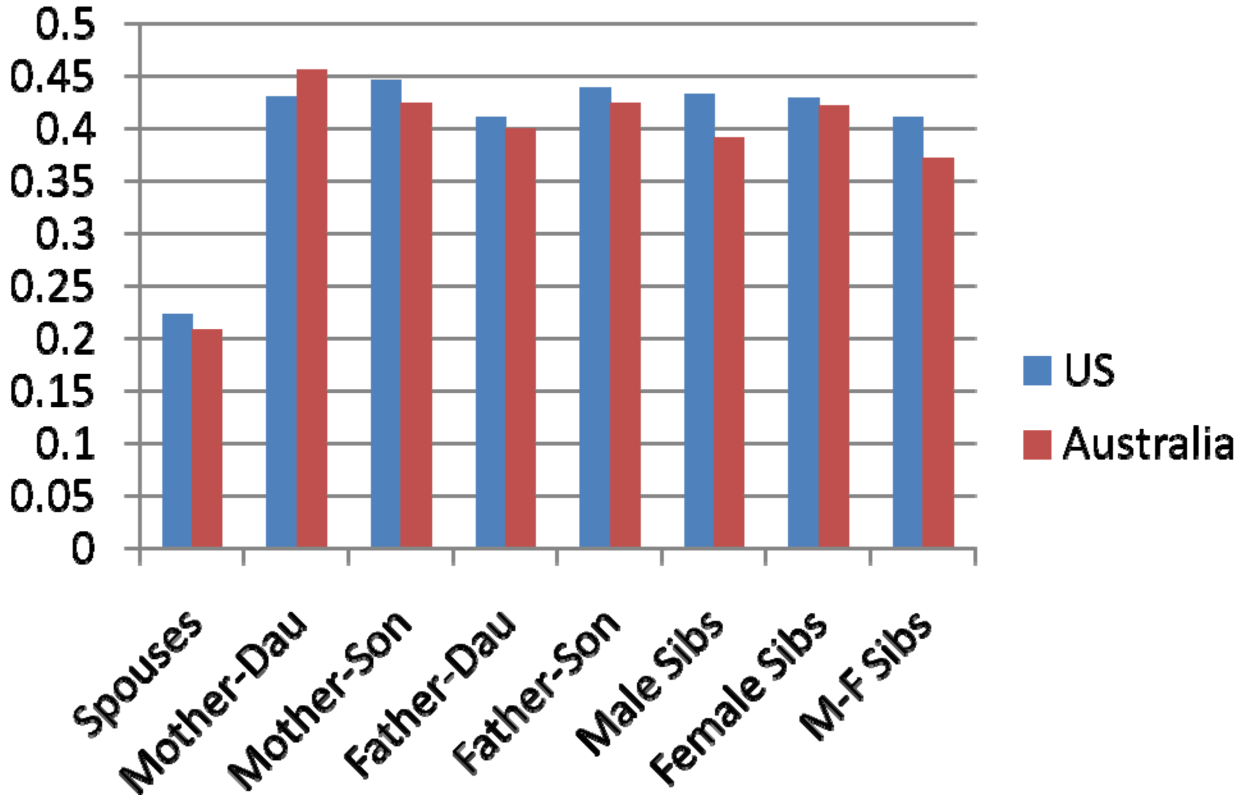
ANZUS 50K: Extended Kinships of Twins



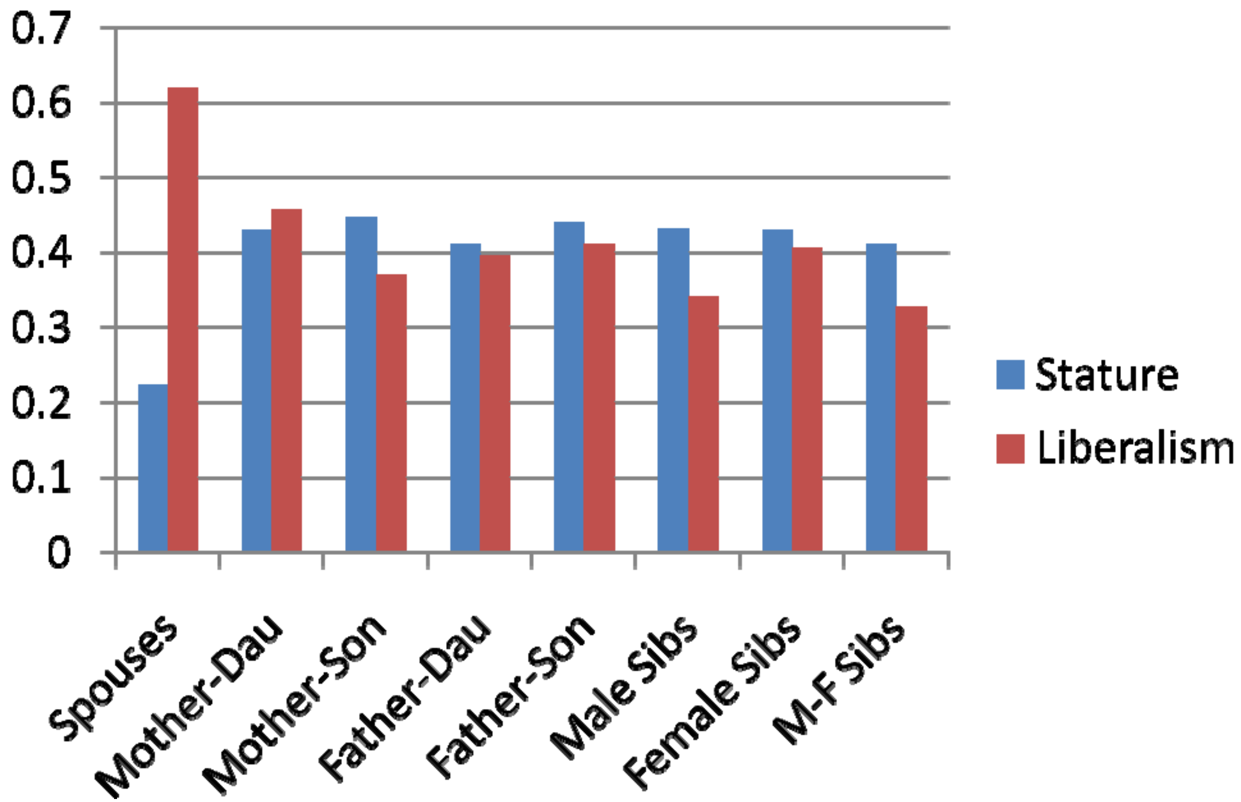
Overall sample sizes

Relationship	# of pairs
Parent-offspring	25018
Siblings	18697
Spouses	8287
DZ Twins	5120
MZ Twins	4623

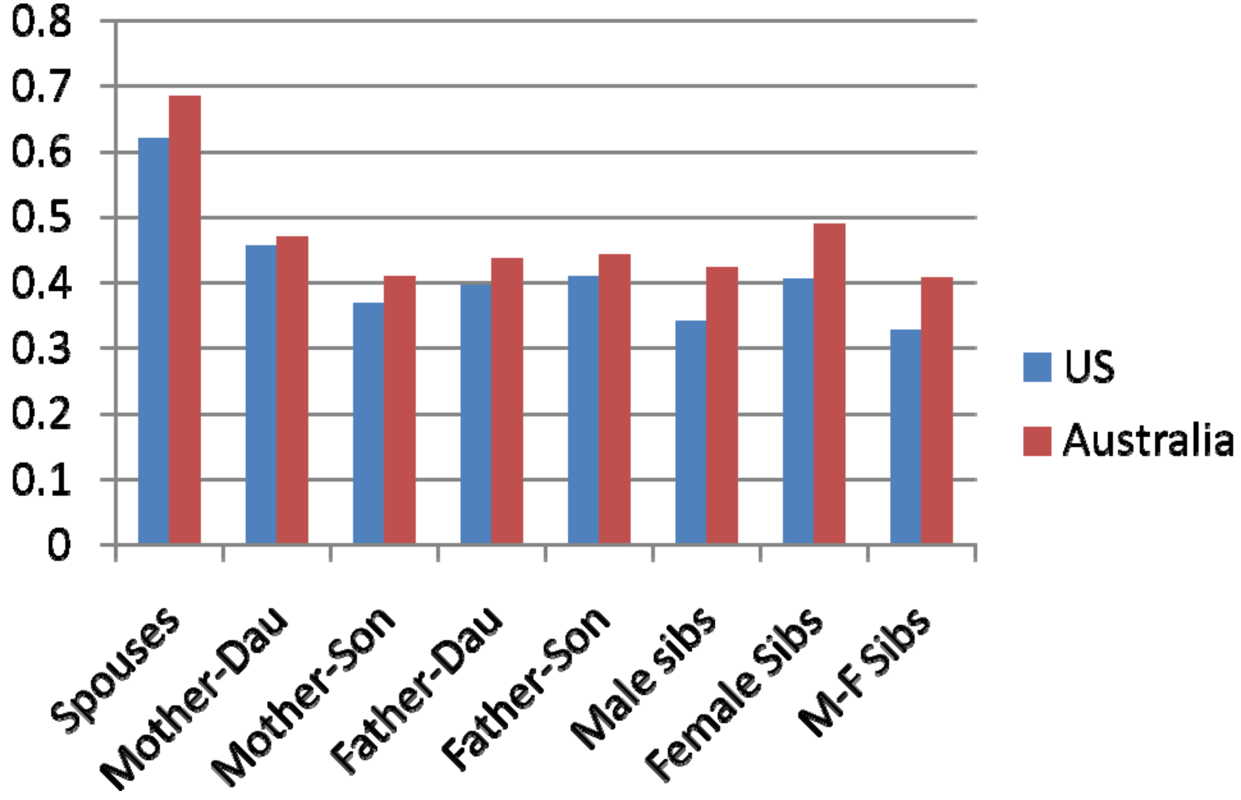
Nuclear Family Correlations for Stature (Virginia 30,000 and OZ 22,000)



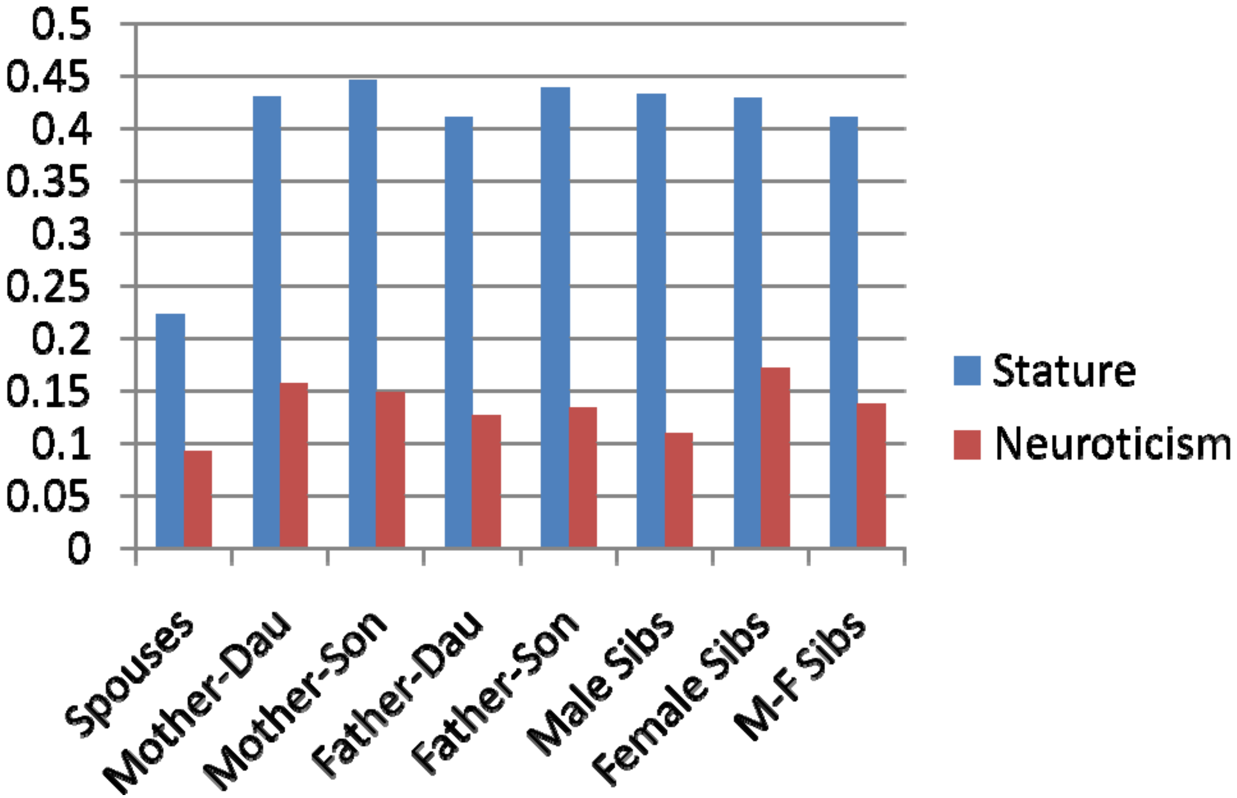
Nuclear Family Correlations for Stature and Liberalism/Conservatism (Virginia 30,000)



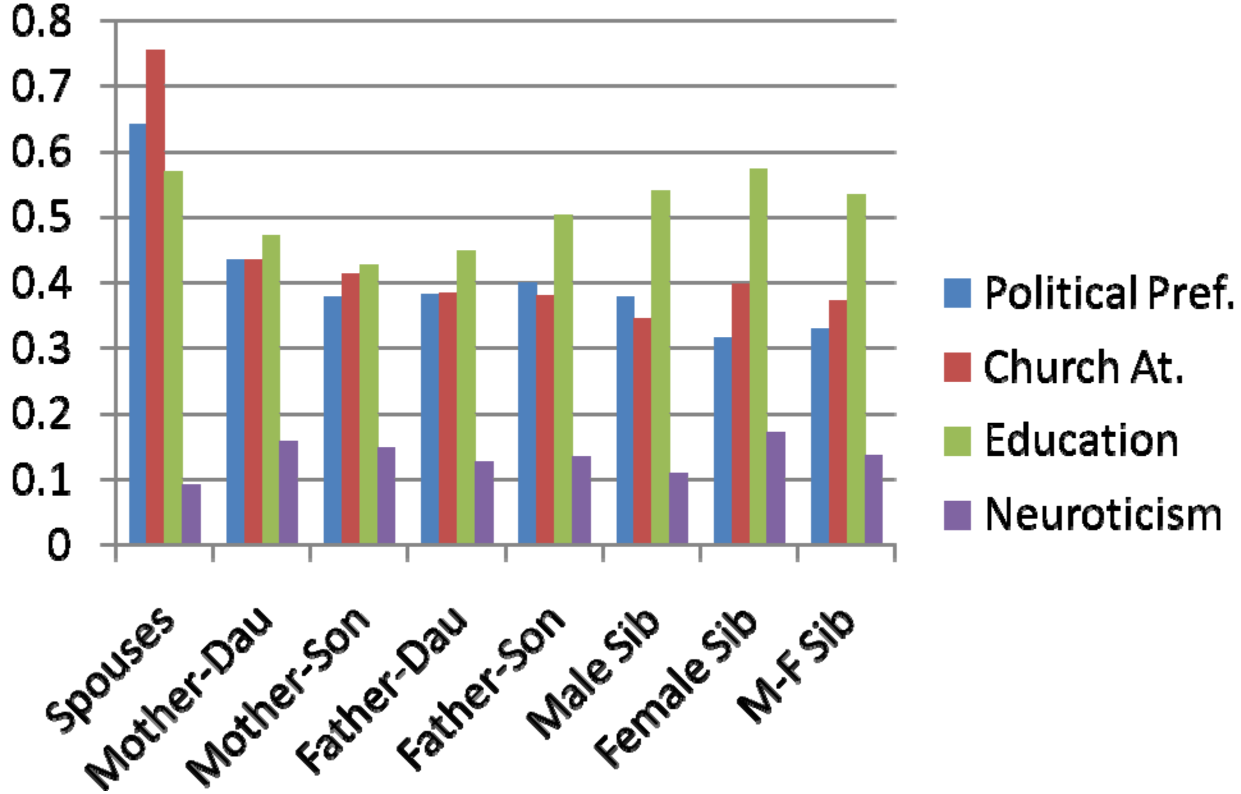
Nuclear Family Correlations for Liberalism/Conservatism (Virginia 30,000 and Australia 22,000)



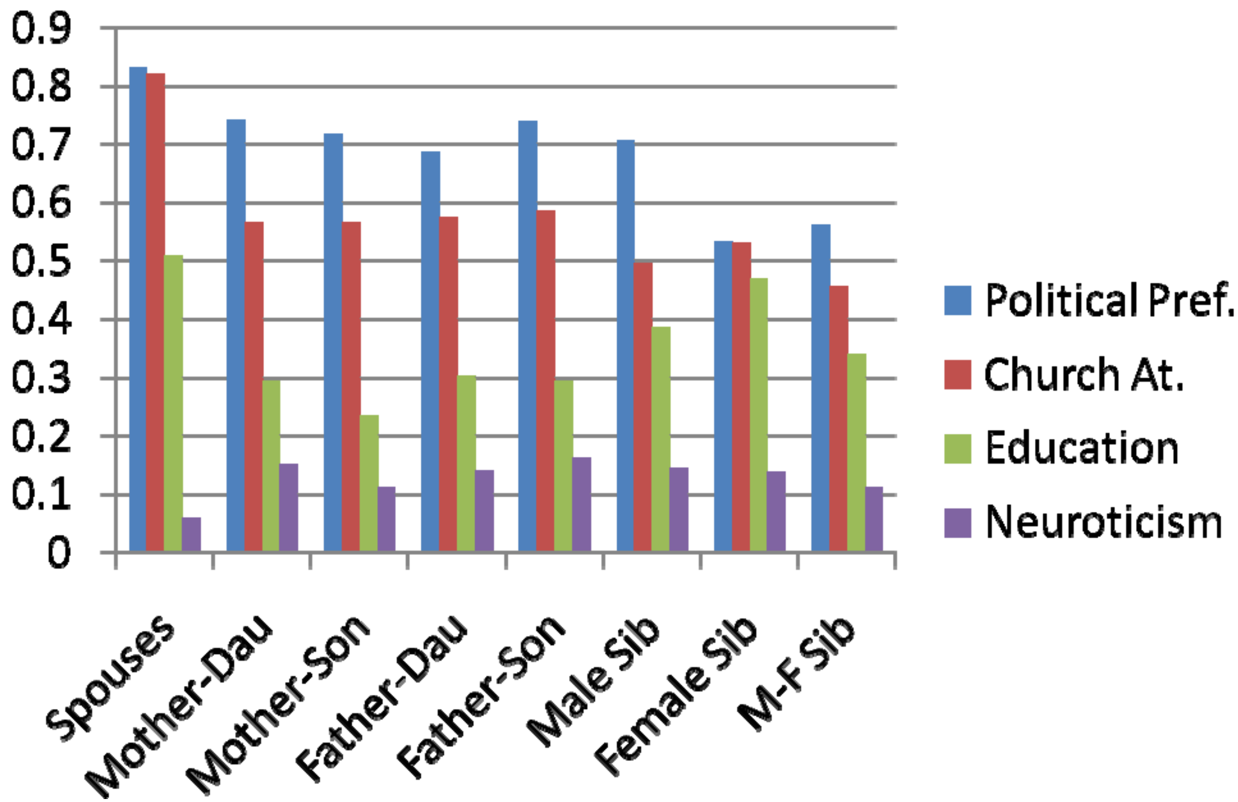
Nuclear Family Correlations for Stature and EPQ Neuroticism (Virginia 30,000)



Nuclear Family Correlations for Socially Significant Variables (Virginia 30,000)



Nuclear Family Correlations for Socially Significant Variables (Australia 22K)



The (Really!) BIG Problem

Families are a mixture of genetic and social factors

A Basic Model

Phenotype=Genotype+Environment

P=G+E {+f(G,E)}

f(G,E) = Genotype-environment interaction and correlation

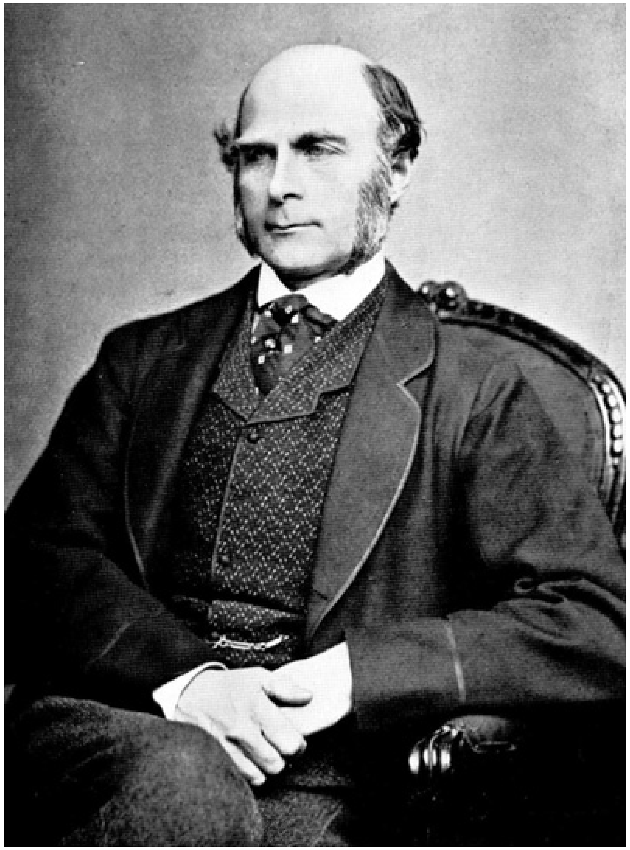
A Basic Model

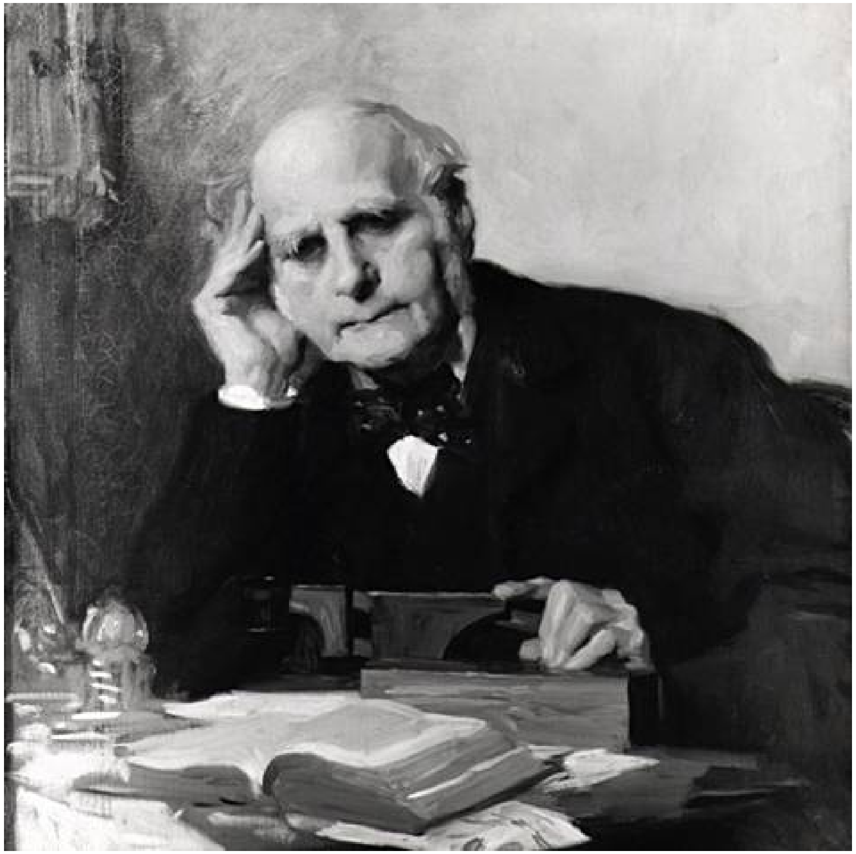
Phenotype=Genotype+Environment

P=G+E {+f(G,E)}

f(G,E) = Genotype-environment interaction and correlation

Francis Galton (1822-1911)





1869: Hereditary Genius

1883: Inquiries into Human Faculty and its Development

1884-5: Anthropometic Laboratory at "National Health Exhibition"

Galton's Solution:

Twins

(Though Augustine may have got there first – 5th cent.)

One (?ideal) solution

Twins separated at birth

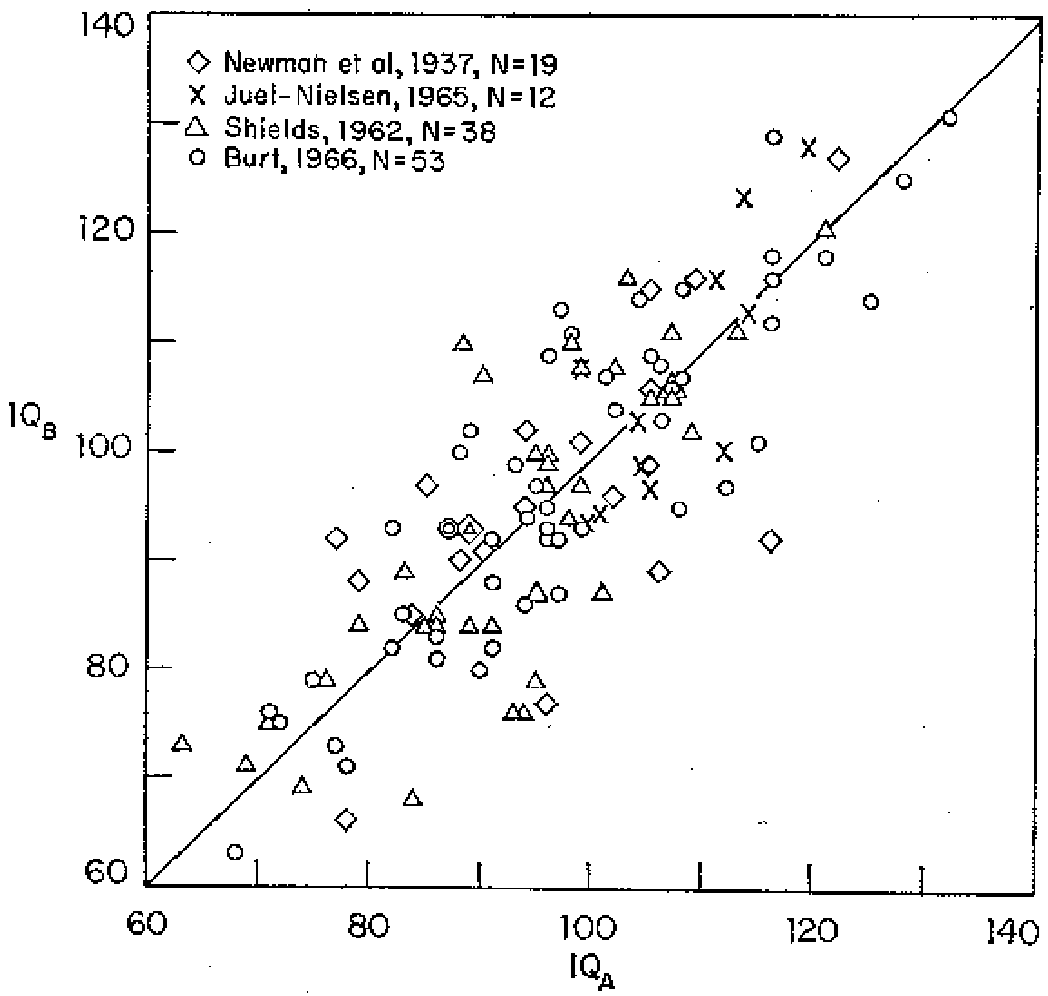


FIGURE 2. *Scatter diagram showing correlation between IQs of 122 sets of co-twins (A and B assigned at random). The obtained intra-class correlation (r_i) is 0·82. The diagonal line represents perfect correlation (r_i = 1·00).*

But separated MZs are rare

An easier alternative:

Identical and non-identical twins reared together:

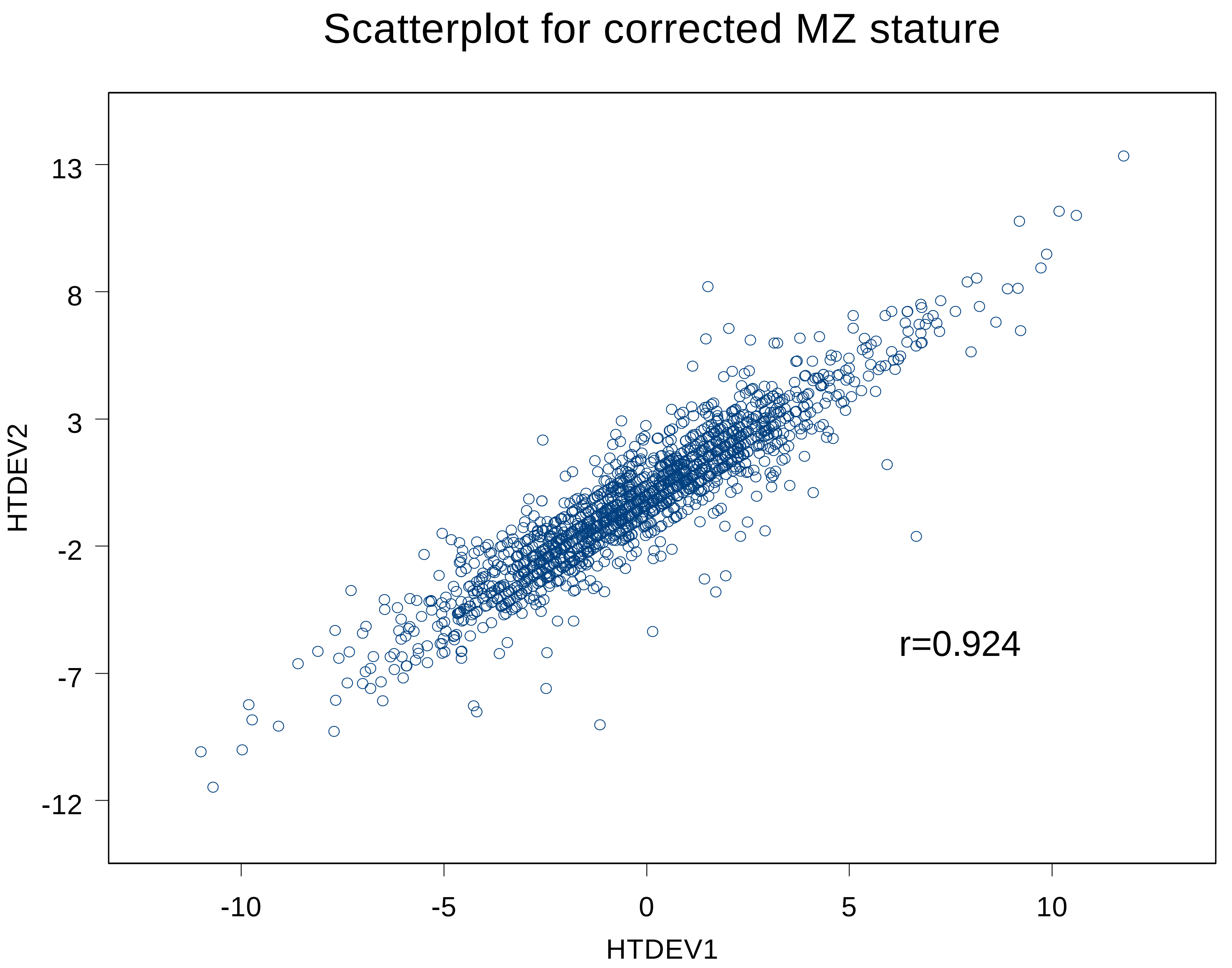
Galton (Again!)

IDENTICAL TWINS

- MONOZYGOTIC: Have IDENTICAL genes (G)
- Come from the same family (C)
- Have unique experiences during life (E)

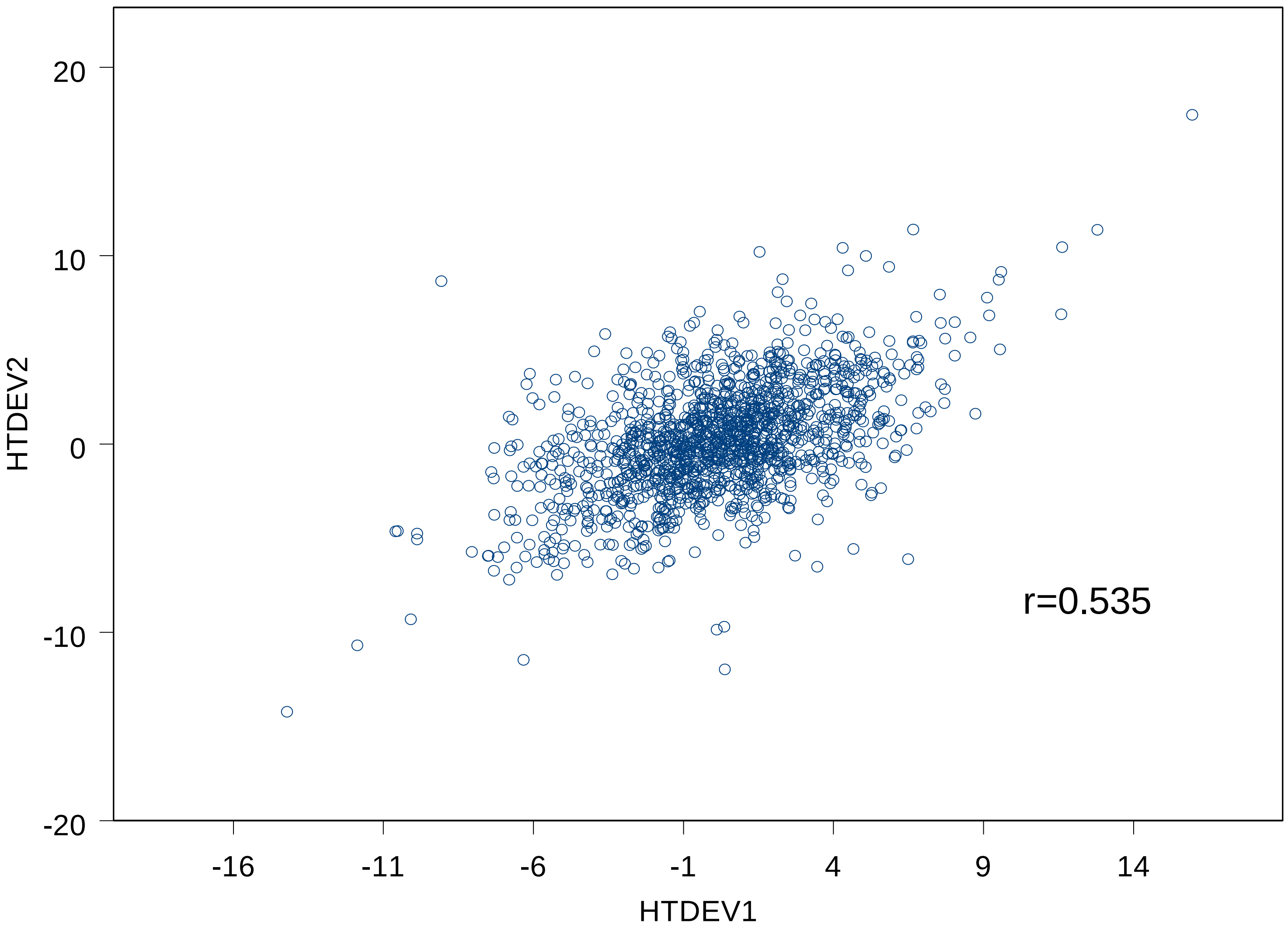
FRATERNAL TWINS

- DIZYGOTIC: Have DIFFERENT genes (G)
- Come from the same family (C)
- Have unique experiences during life (E)



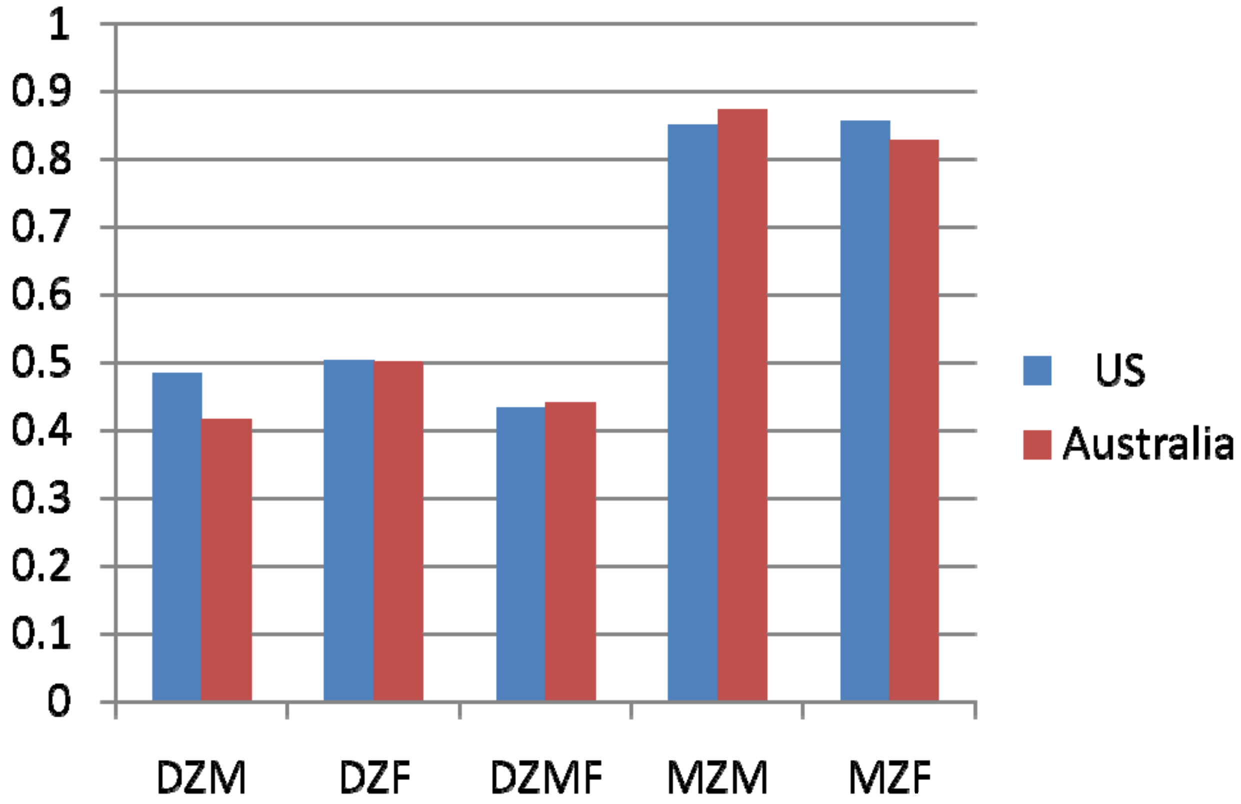
Data from the Virginia Twin Study of Adolescent Behavioral Development

Scatterplot for age and sex corrected stature in DZ twins

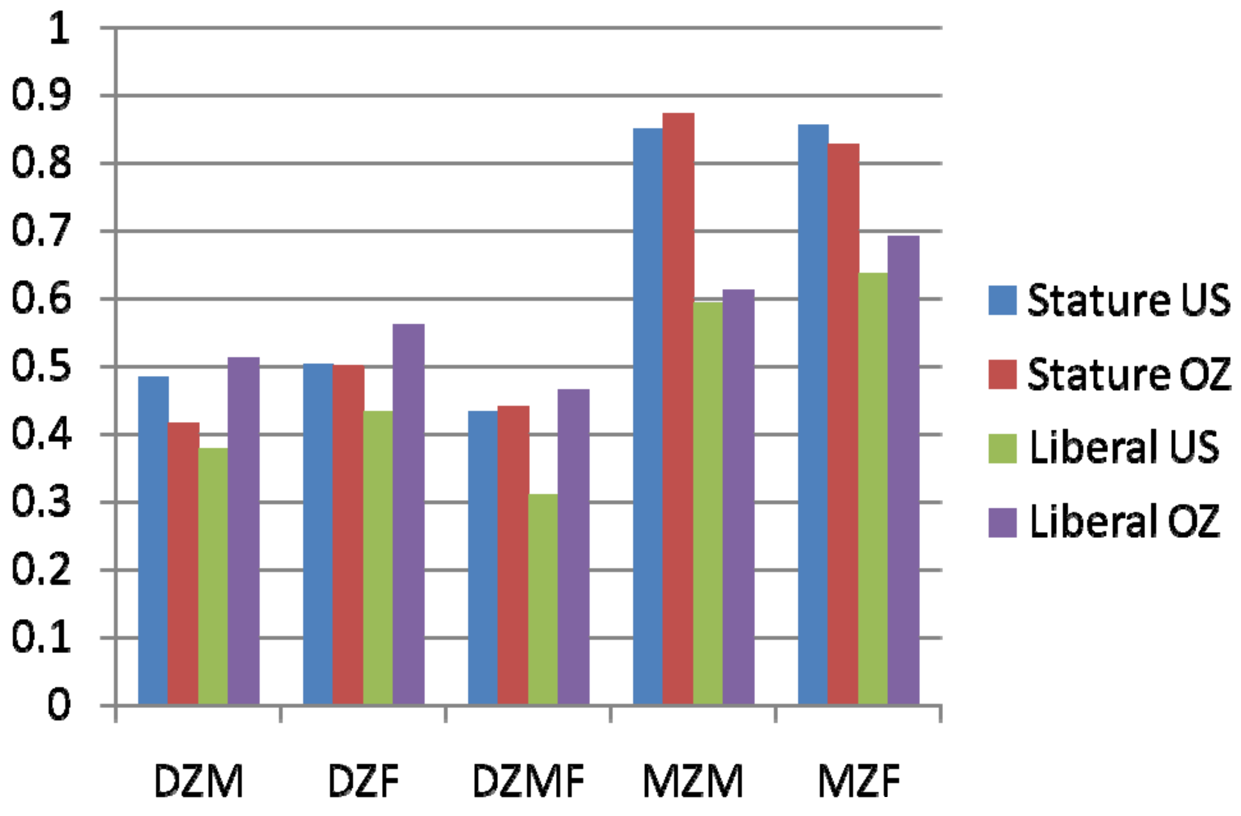


Data from the Virginia Twin Study of Adolescent Behavioral Development

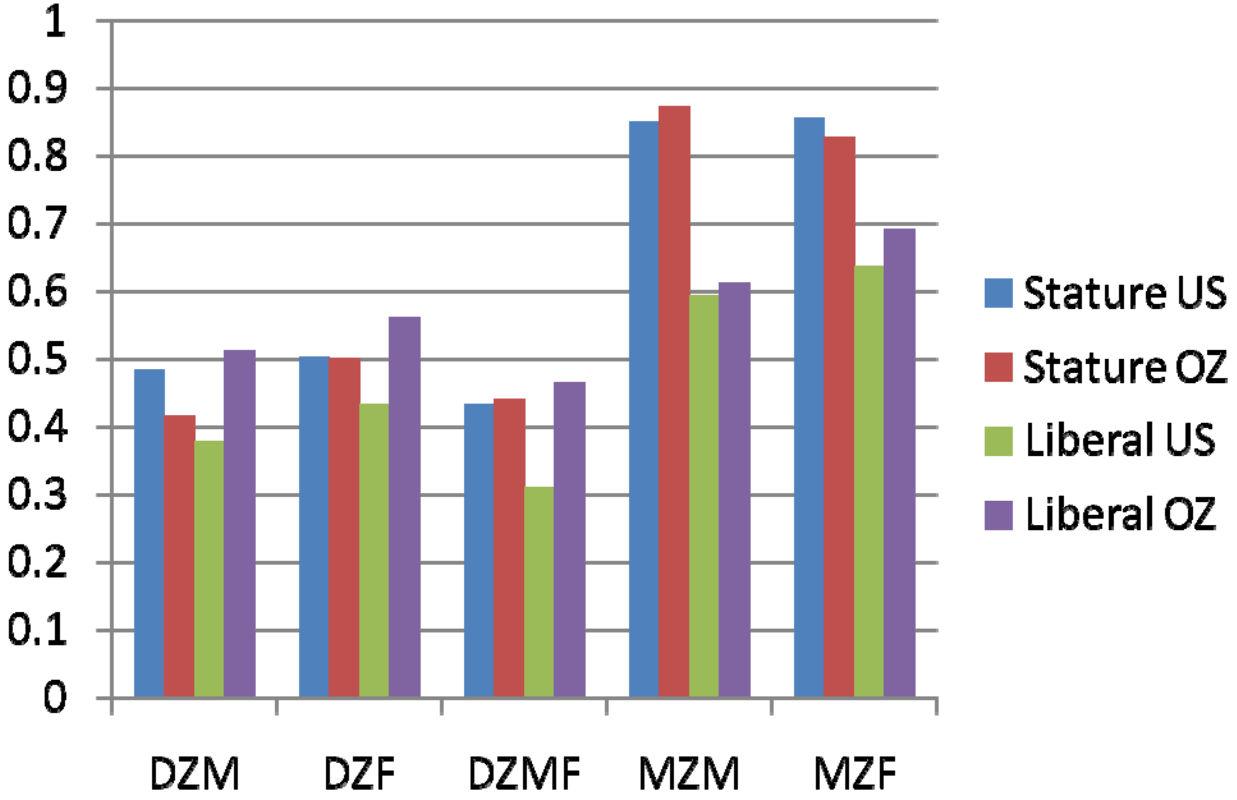
Twin Correlations for Adult Stature
(Virginia 30,000 and Australia 22,000)



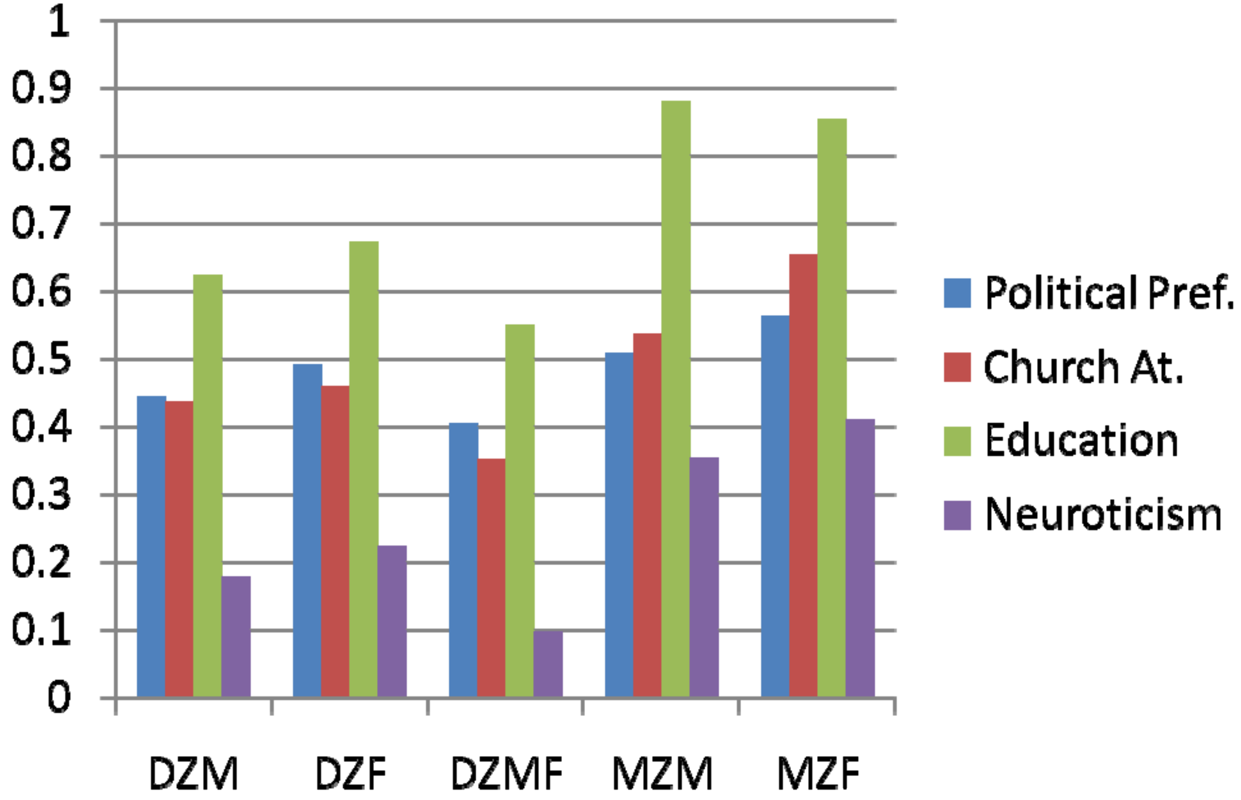
Twin Correlations for Stature and Liberalism (Virginia 30,000 and Australia 22,000)



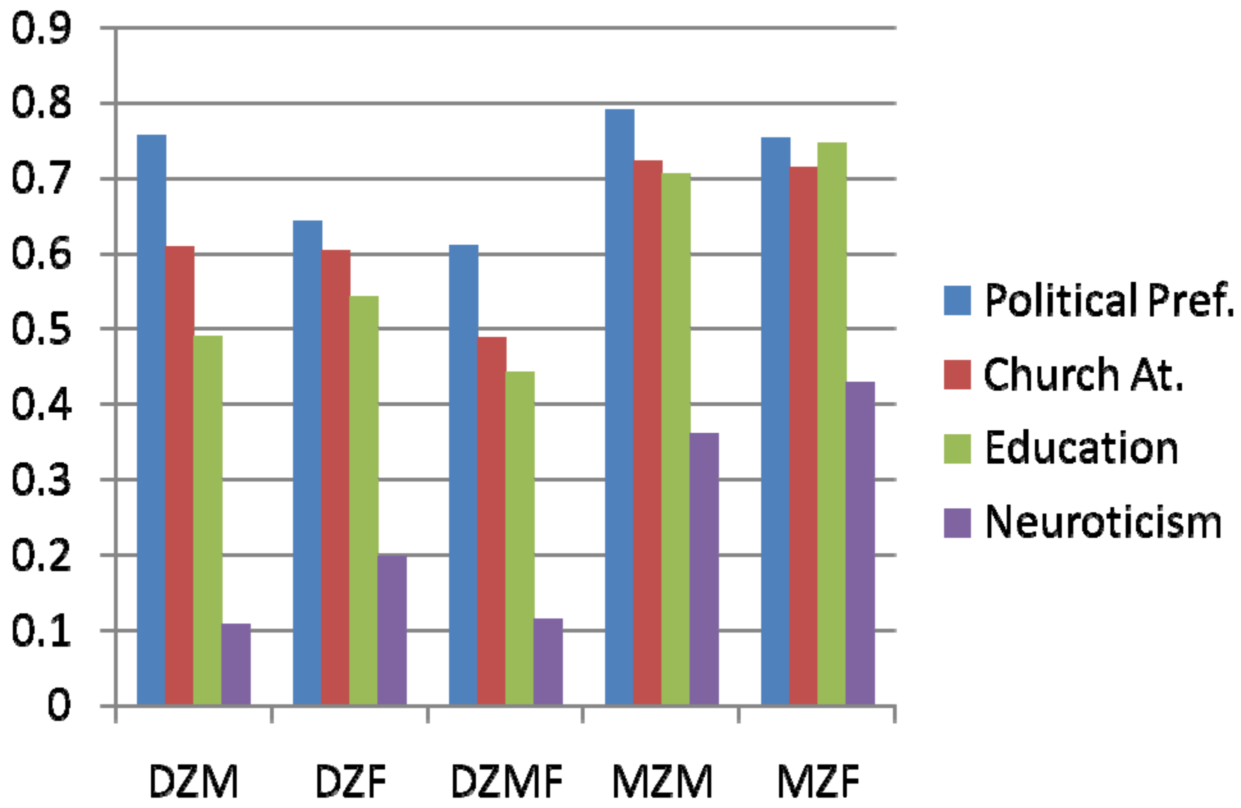
Twin Correlations for Stature and Liberalism (Virginia 30,000 and Australia 22,000)



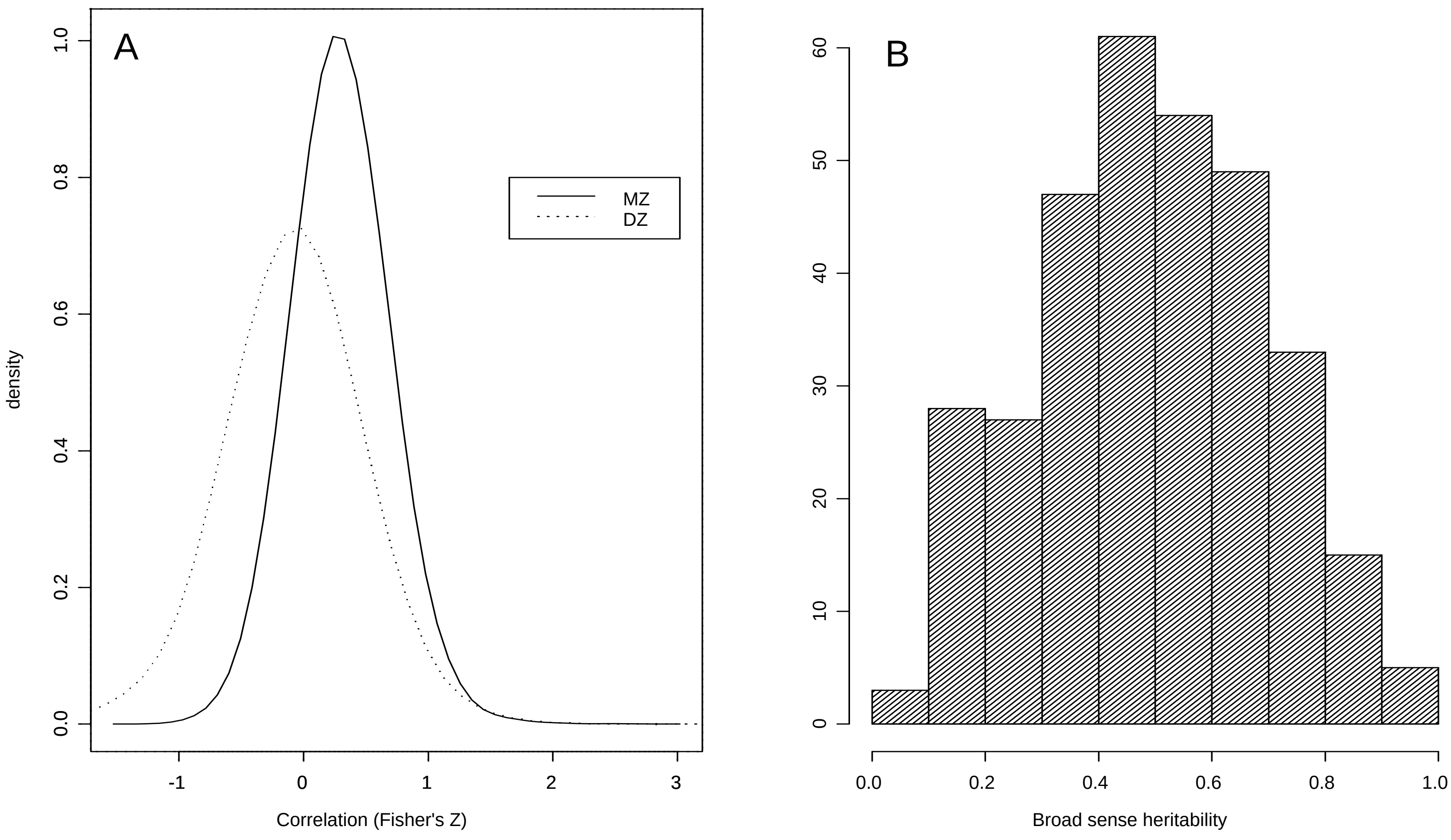
Twin Correlations for Socially Significant Variables (Virginia 30,000)



Twin Correlations for Socially Significant Variables (Australia 22,000)



Twin correlations for gene expression



York et al.

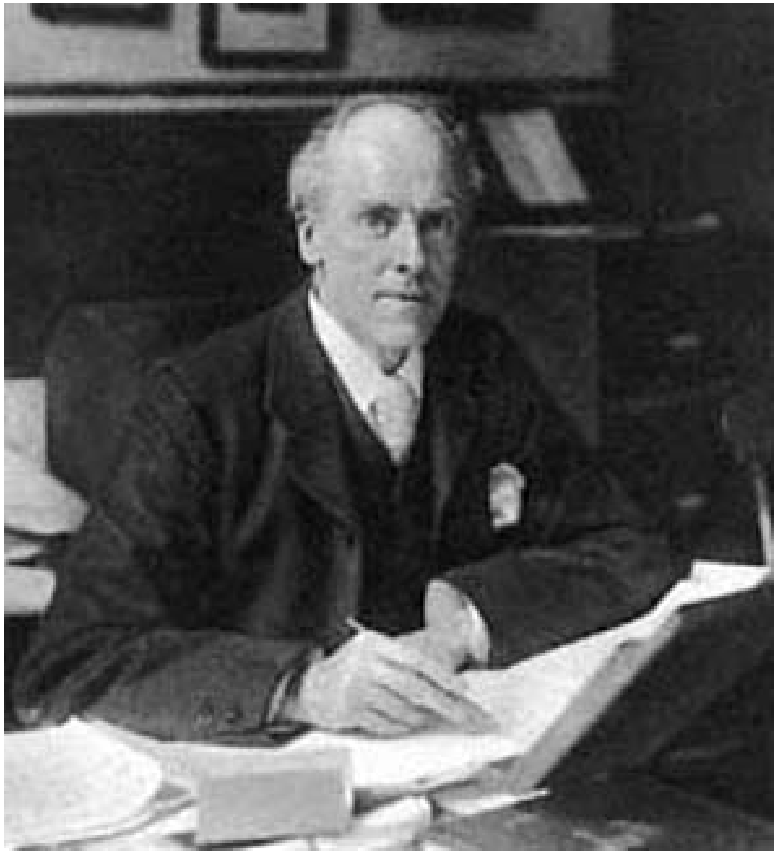
“Quantitative Genetics”

Analysis of the patterns and mechanisms underlying variation in continuous traits to resolve and identify their genetic and environmental causes.

Gregor Mendel (1822-1884)

1865: “Experiments in Plant Hybridization”

Karl Pearson (1857-1936)



1903: On the Laws of Inheritance in Man: I Physical Characteristics (with Alice Lee)

1904: II Mental and Moral Characteristics

1914: The Life, Letters and Labours of Francis Galton

“Mendelian” Crosses with Quantitative Traits

Mendelian Basis of Continuous Variation? Experimental Breeding Experiments

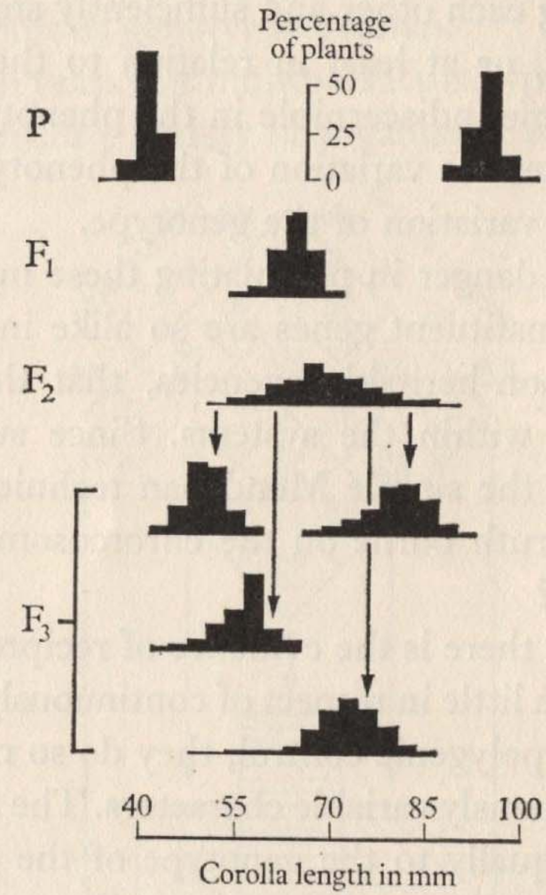


Figure 2 The inheritance of corolla length in *Nicotiana longiflora* (East 1915).

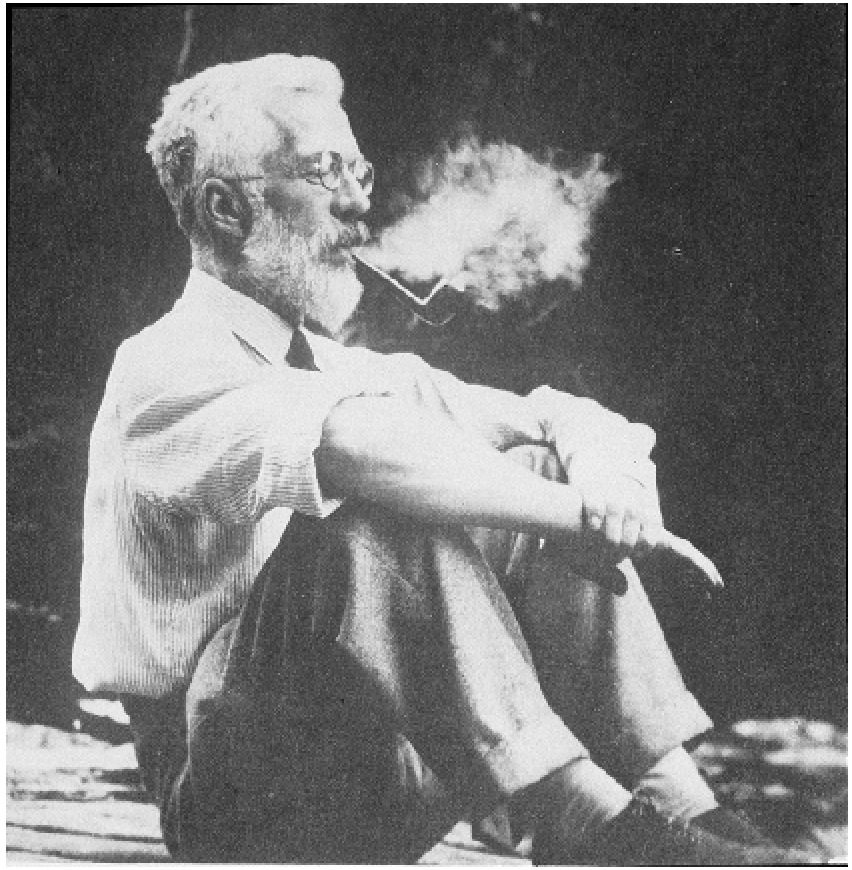
Experiments Show:

- Variation within inbred lines: Environment
- F_1's typically show same within-line variation
- F_2's more variable: Mirrors Mendelian segregation of Mendel's classical hybridization experiments
- Average differences between individual F_2 plants continue to progeny generations (F_3's etc.)

Description of East's Experiment

Figure 2 The inheritance of corolla length in *Nicotiana longiflora* (East 1915). For ease of presentation, the results are shown as the percentage frequencies with which individuals fall into classes, each covering a range of 3 mm in corolla length and centred on 34, 37, 40, etc., mm. This grouping is quite artificial and the apparent discontinuities spurious: corolla length actually varies continuously. The means of F_1 and F_2 are intermediate between those of the parents. The means of the four F_3 families are correlated with the corolla length of the F_2 plants from which they came, as indicated by the arrows. Variation in parents and F_1 is all non-heritable, and hence is less than that in F_2 which shows additional variation arising from the segregation of the genes concerned in the cross. Variation in F_3 is on the average less than that of F_2 but greater than that of parents and F_1. Its magnitude varies among the different F_3's, according to the number of genes which are segregating.

Ronald Fisher (1890-1962)



1918: On the Correlation Between Relatives on the Supposition of Mendelian Inheritance
1921: Introduced concept of “likelihood”
1930: The Genetical Theory of Natural Selection
1935: The Design of Experiments

Fisher developed mathematical theory that reconciled Mendel's work with Galton and Pearson's correlations

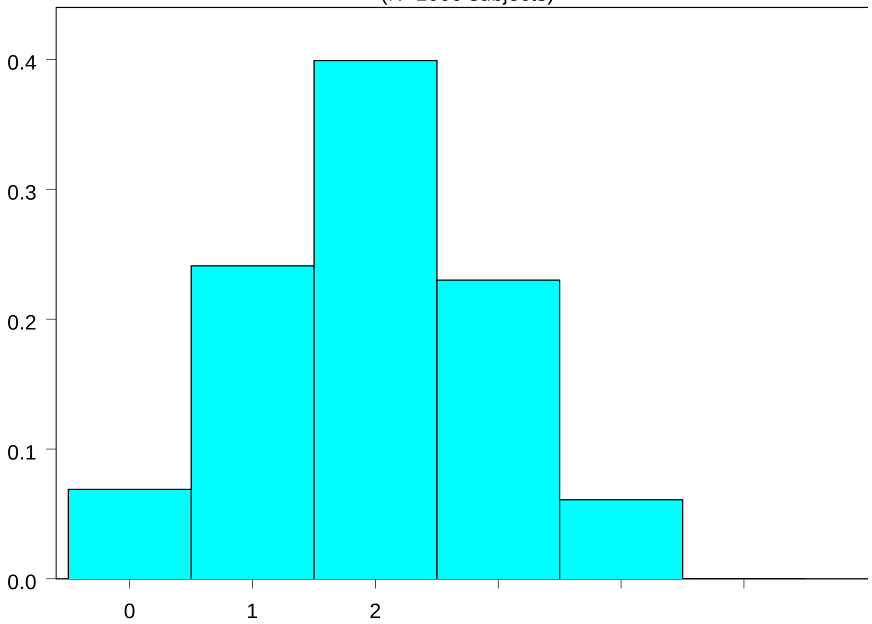
XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. By R. A. Fisher, B.A. *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.

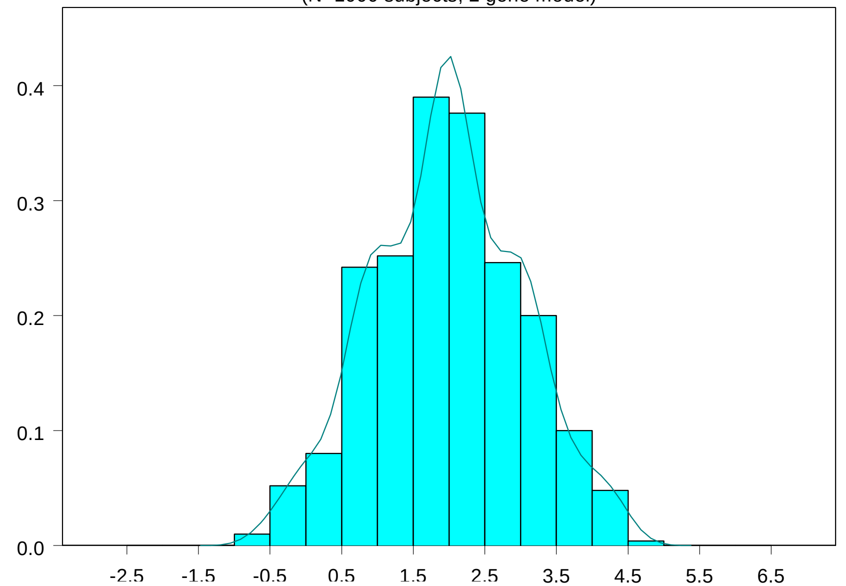
a. Distribution of scores produced by two genes

(N=1000 subjects)



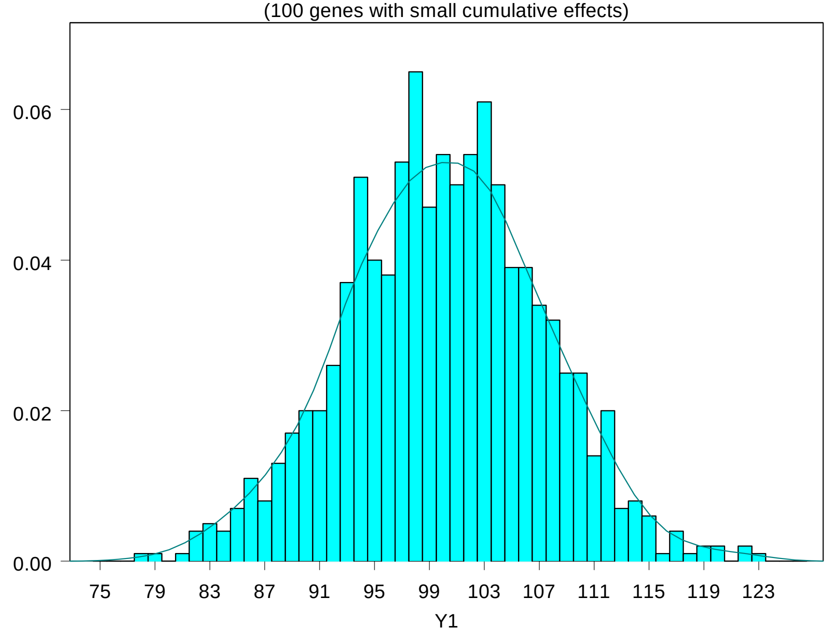
b. The "smoothing" effect of the environment

(N=1000 subjects, 2 gene model)



c. Continuous distribution of polygenic trait

(100 genes with small cumulative effects)



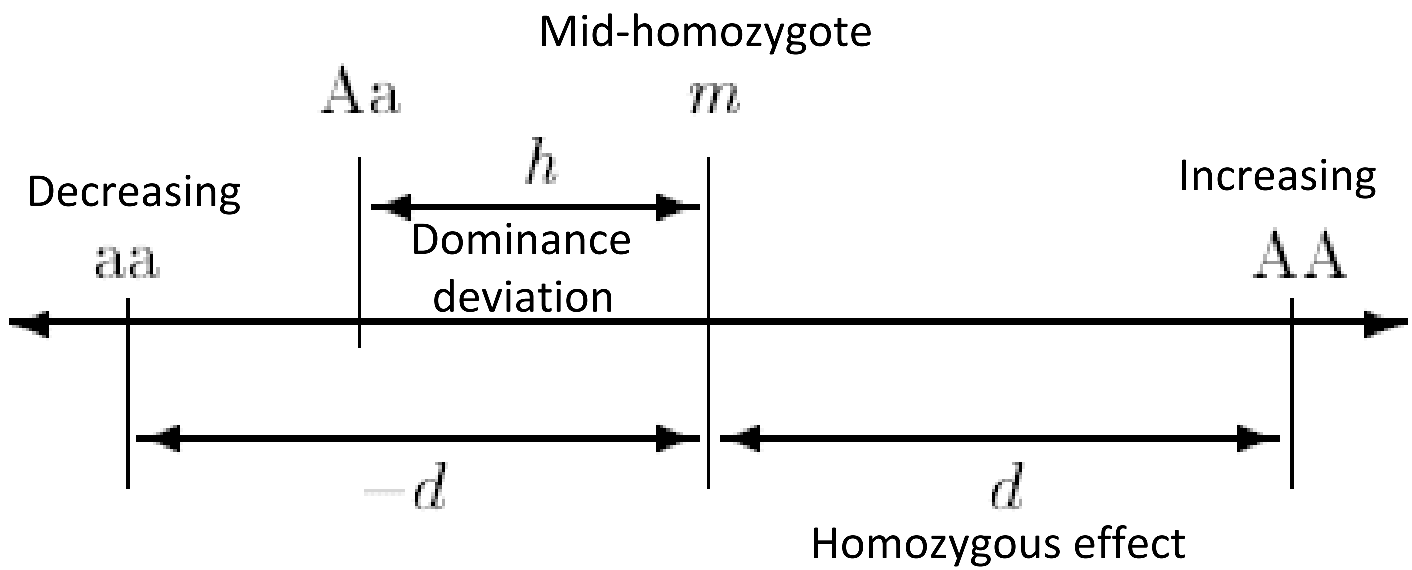
Fisher (1918): Basic Ideas

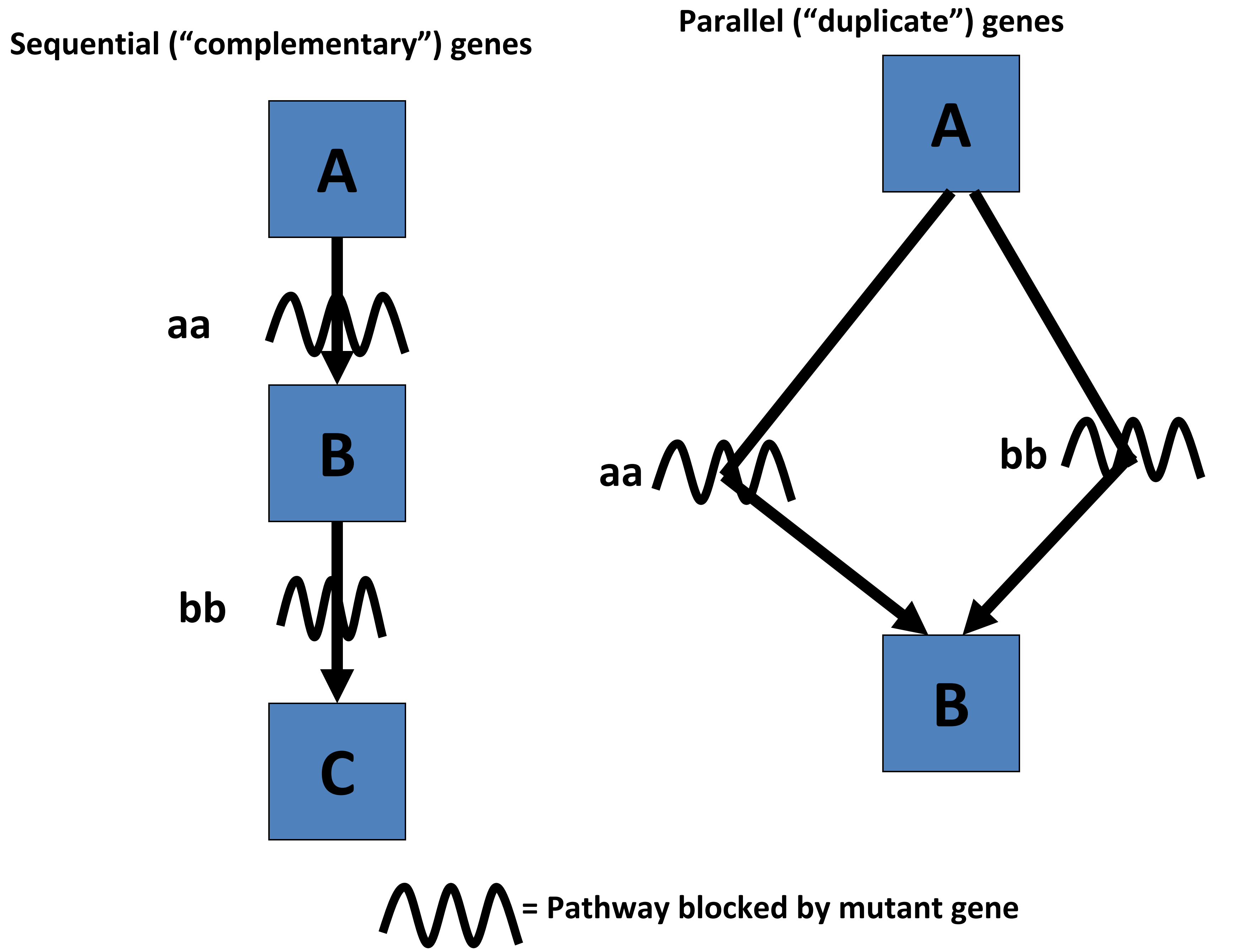
- Continuous variation caused by lots of genes ("polygenic inheritance")
- Each gene followed Mendel's laws
- Environment smoothed out genetic differences
- Genes may show different degrees of "dominance"
- Genes may have many forms ("mutliple alleles")
- Mating may not be random ("assortative mating")
- Showed that correlations obtained by e.g. Pearson and Lee were explained well by polygenic inheritance

Kenneth Mather 1911-1990

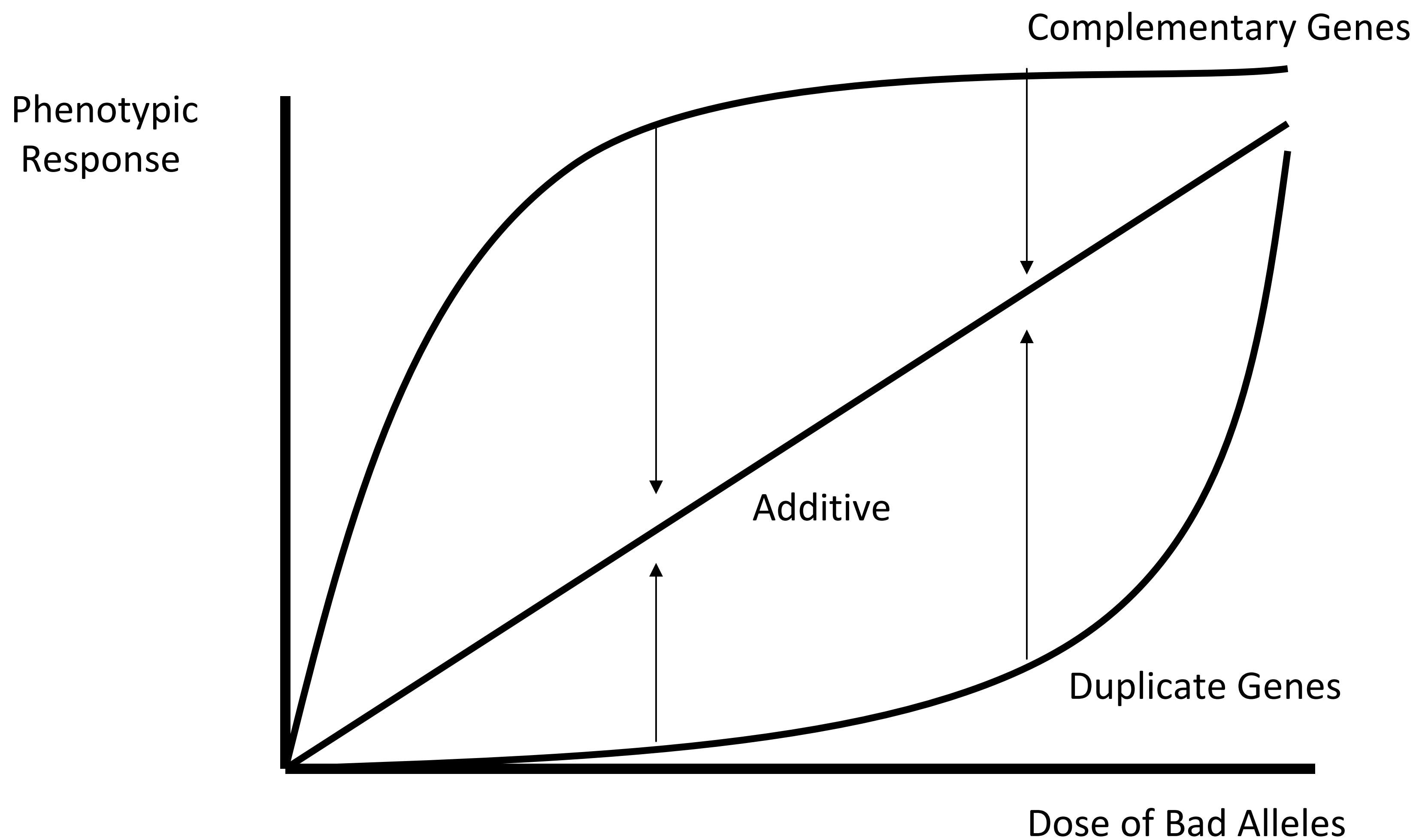
John Jinks 1929-1987

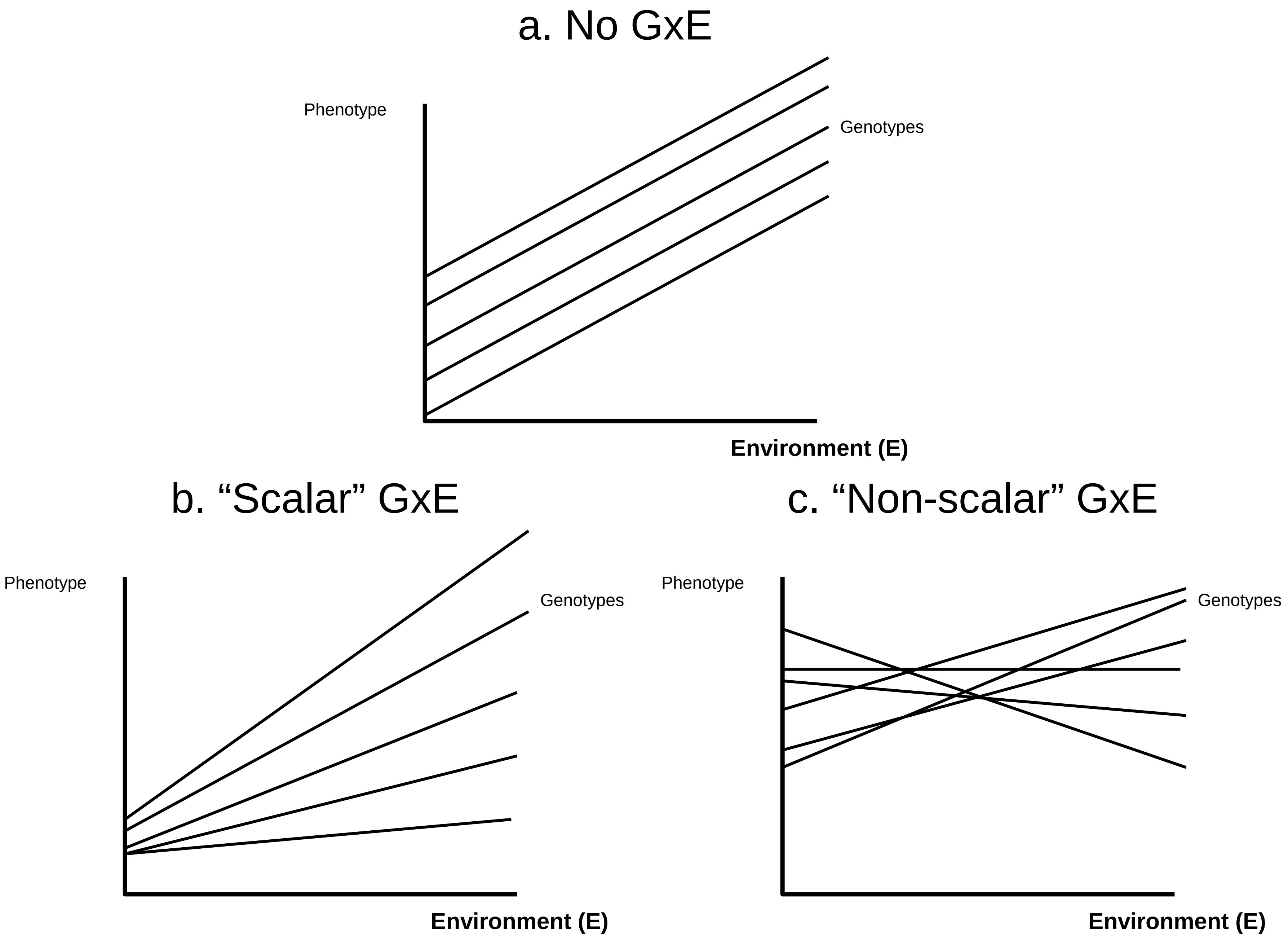
Basic Model for Effects of a Single Gene on a Quantitative Trait



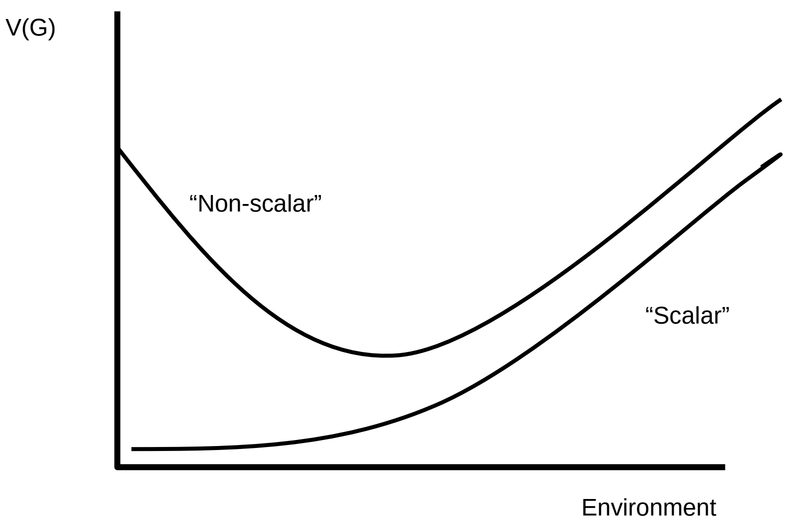

Sequential ("complementary") genes
A
aa
B
bb
C
Parallel ("duplicate") genes
A
aa
bb
B
= Pathway blocked by mutant gene

Combining pathways

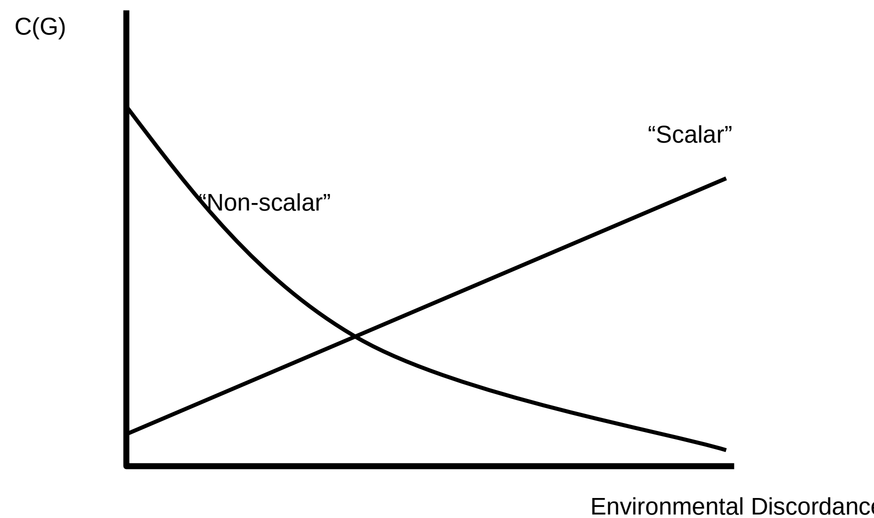



a. No GxE
Phenotype
Genotypes
Environment (E)
b. “Scalar” GxE
Phenotype
Genotypes
Environment (E)
c. “Non-scalar” GxE
Phenotype
Genotypes
Environment (E)

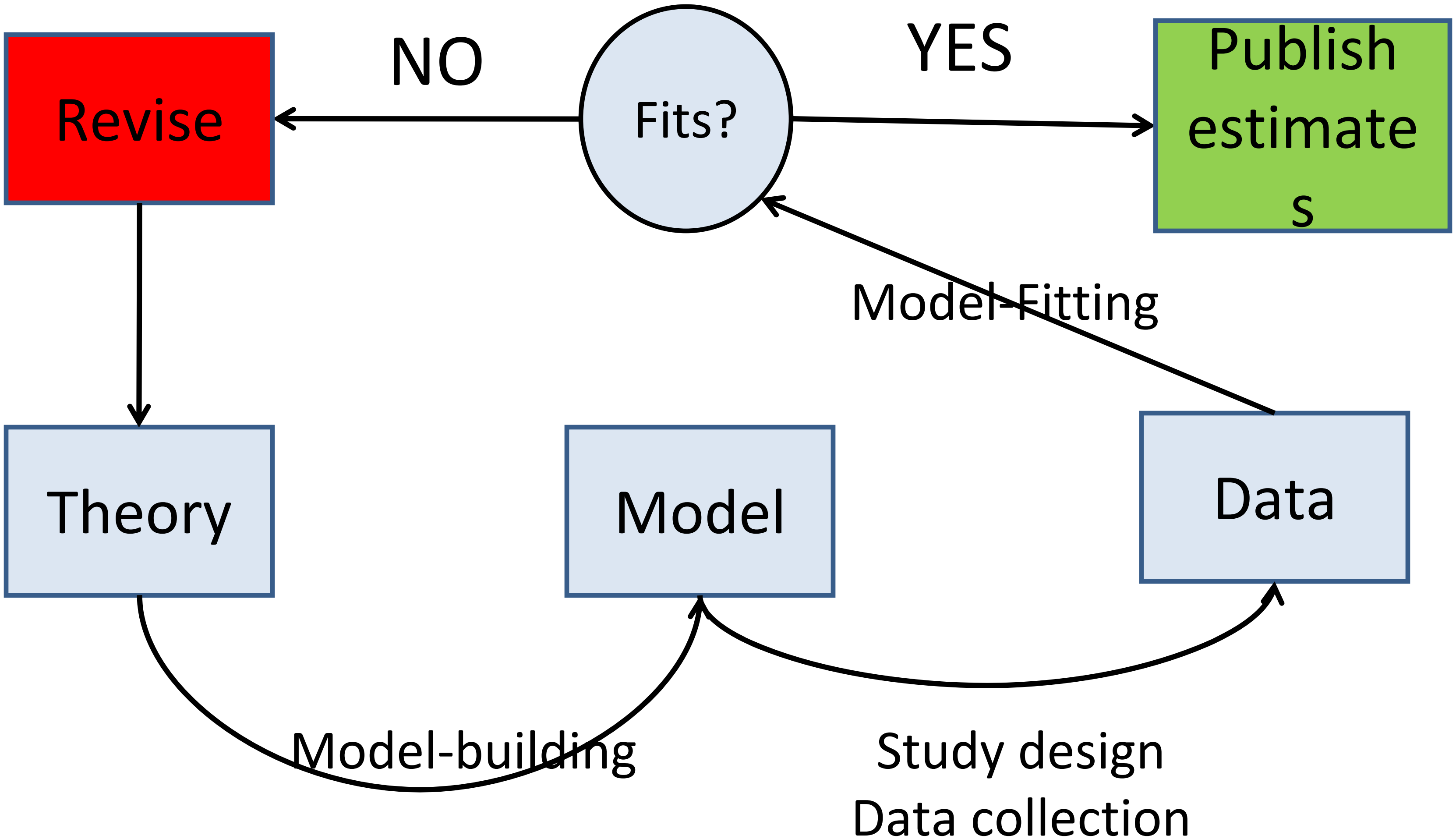
a. Genetic Variance Under GxE



b. Genetic Covariance Under GxE



"The Logic of Scientific Discovery"



Statisical approach

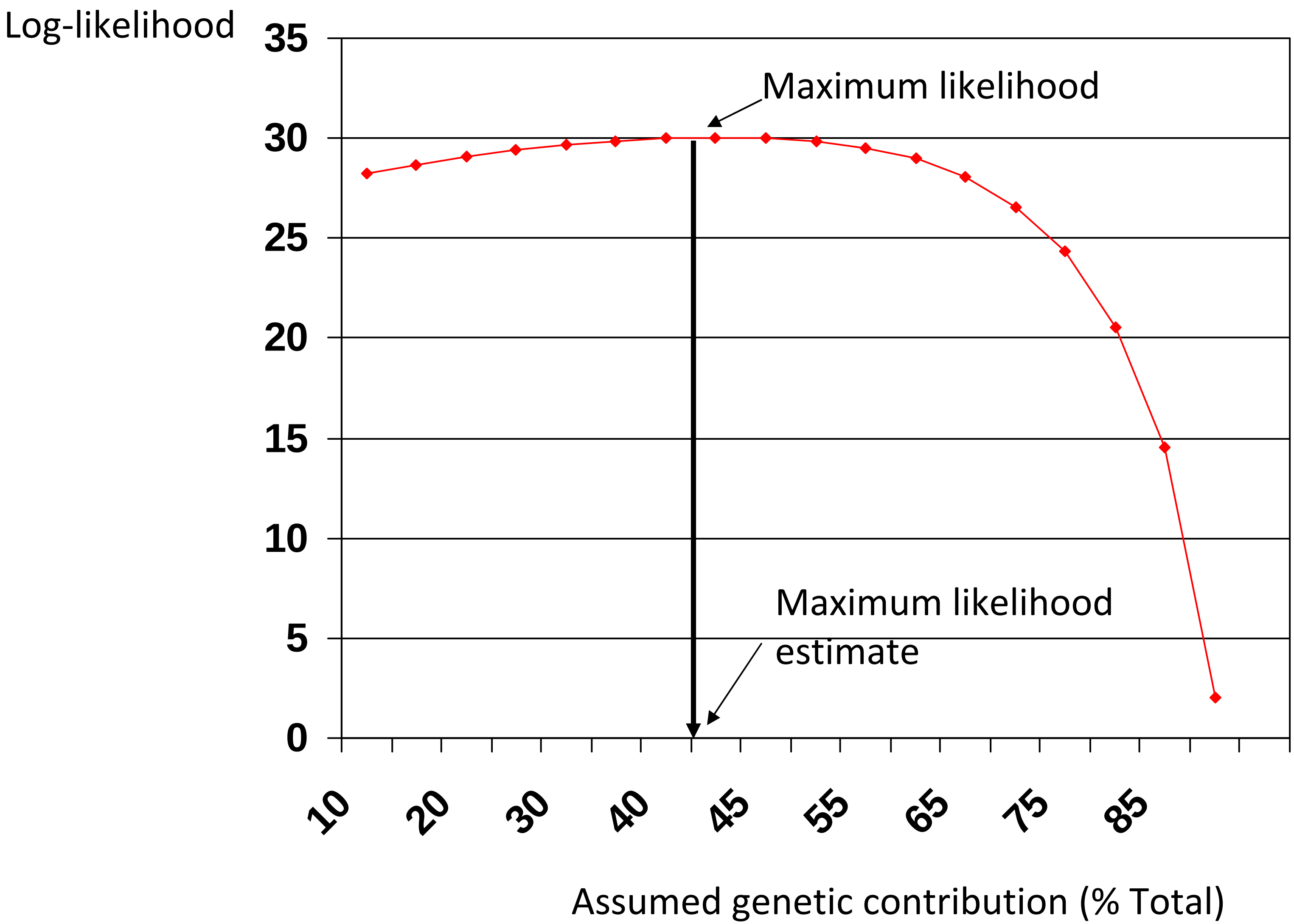
"Likelihood" (Fisher)

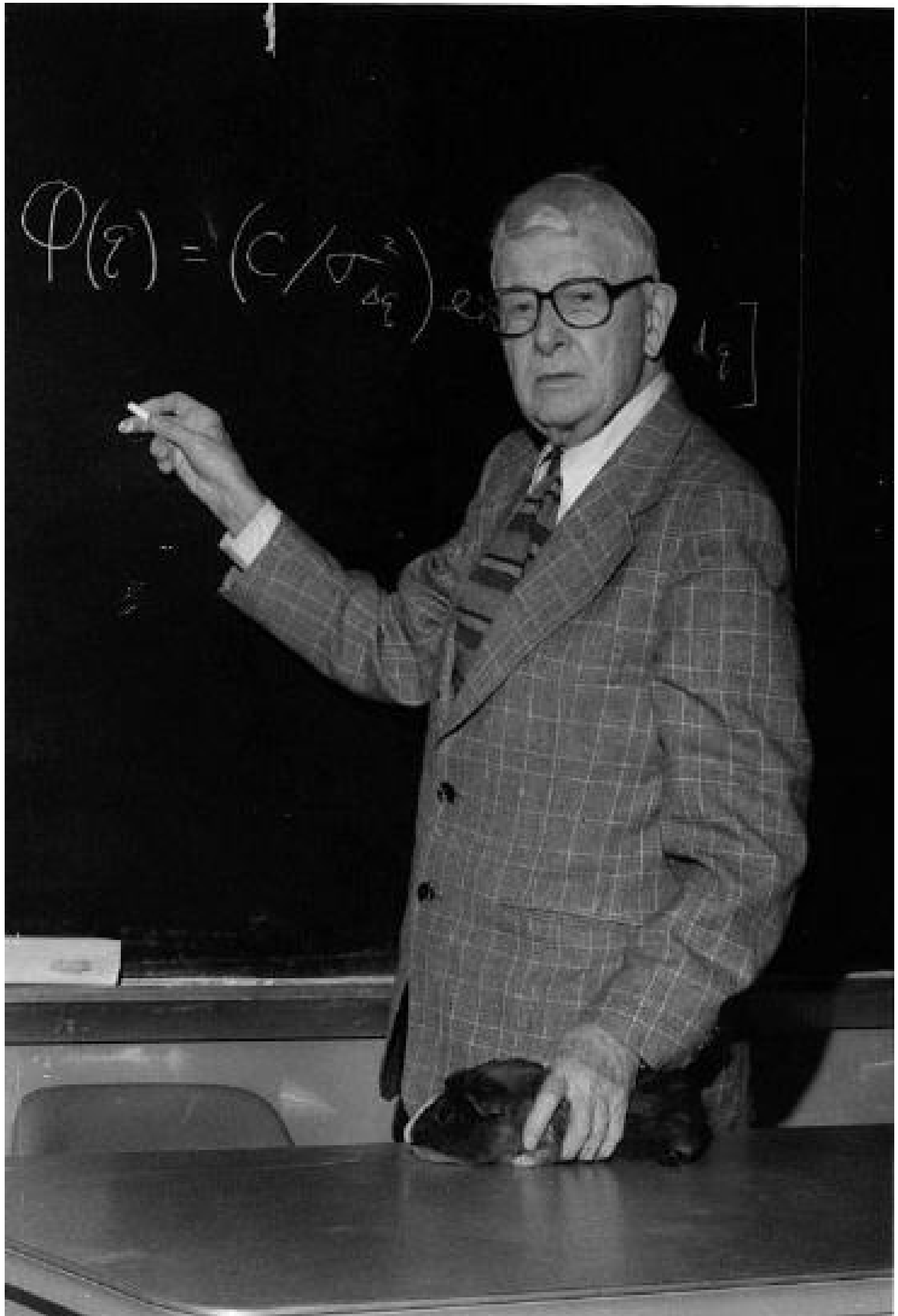
Some models and values of quantities ("parameters", V_A, V_D etc) are "unlikely" to produce the data.

Choose those parameters values for that make the data "most likely", i.e. maximum likelihood.

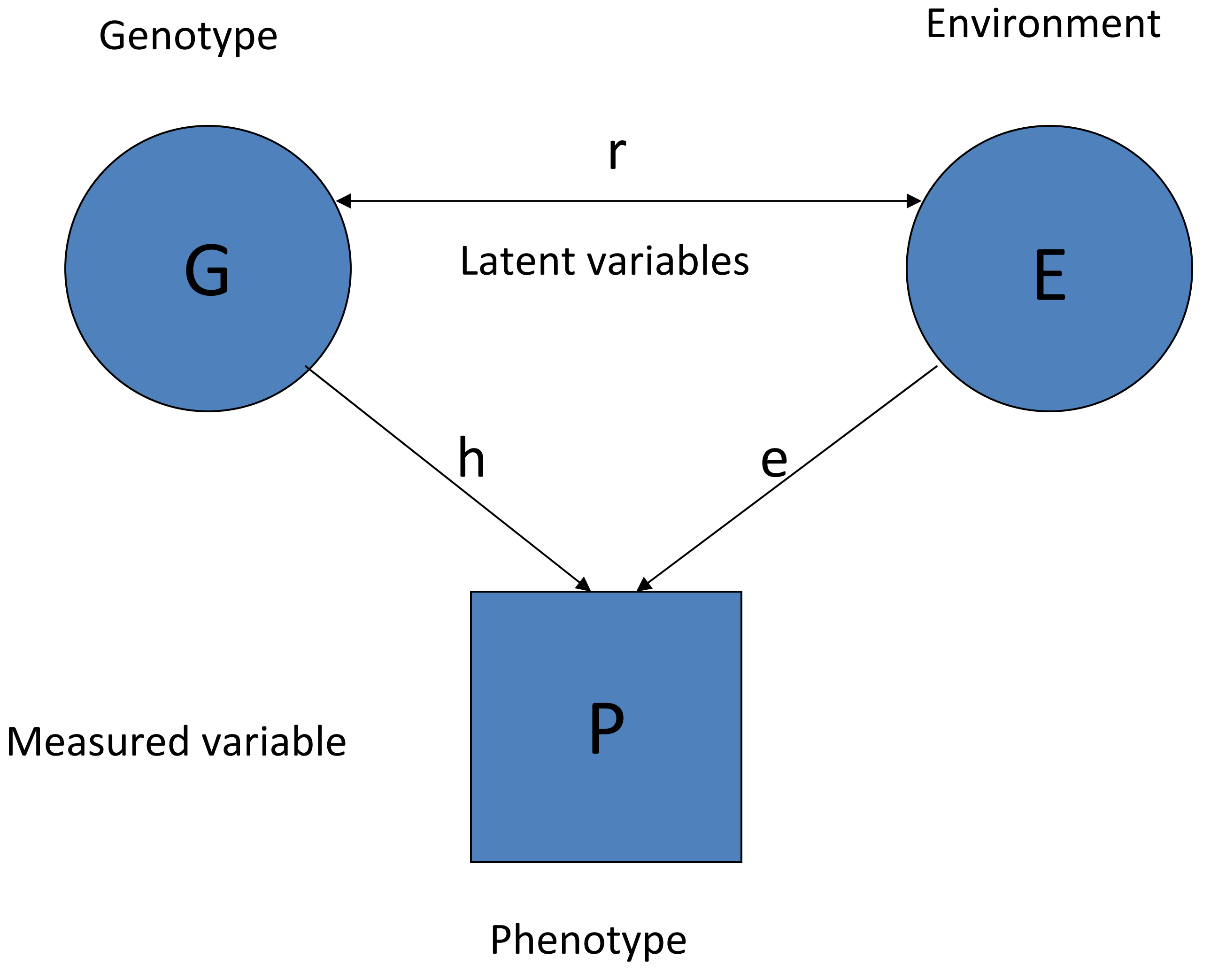
General statistical approach: applied widely in genetics

Example: Log-likelihood of twin correlations for different genetic contributions

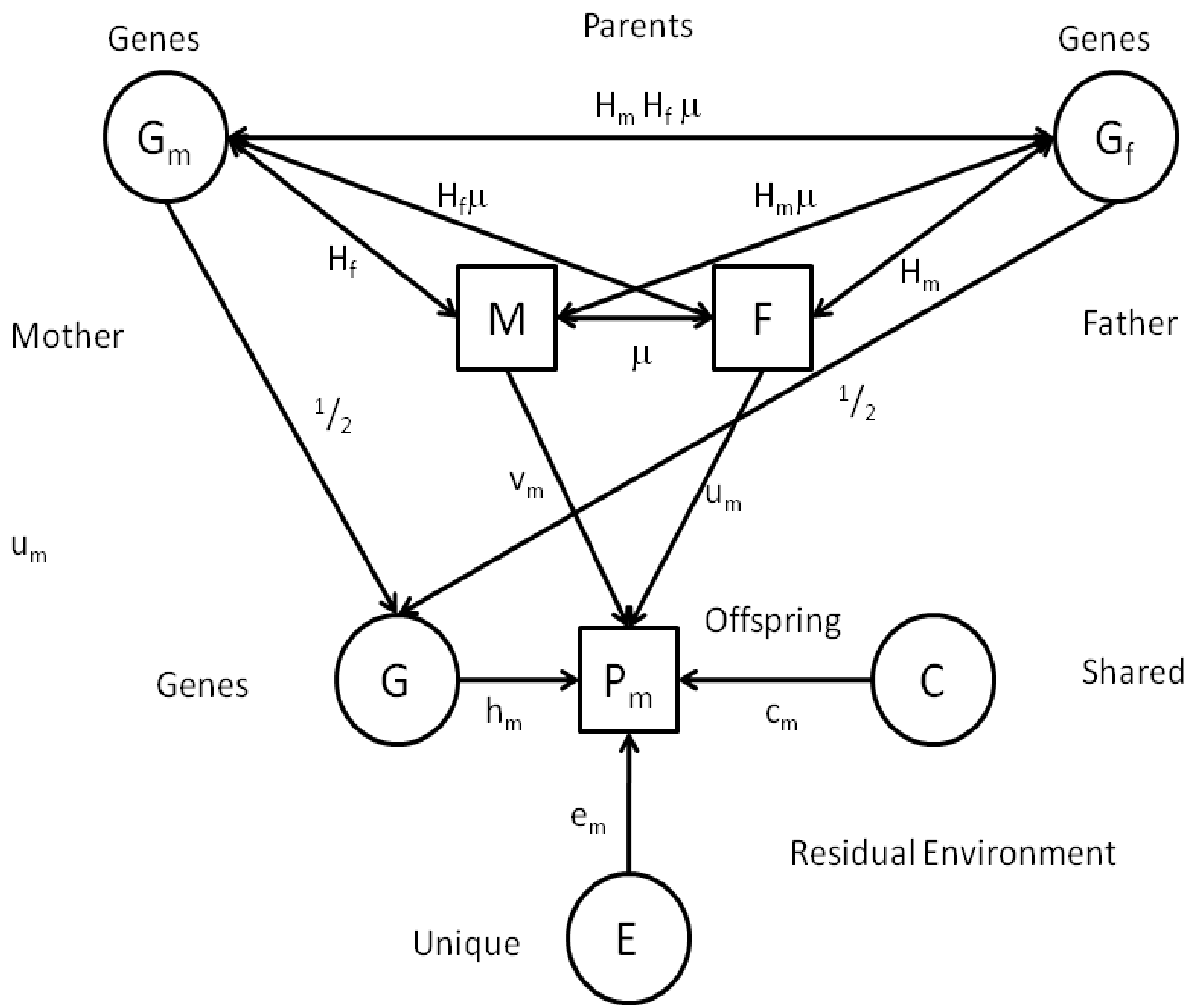




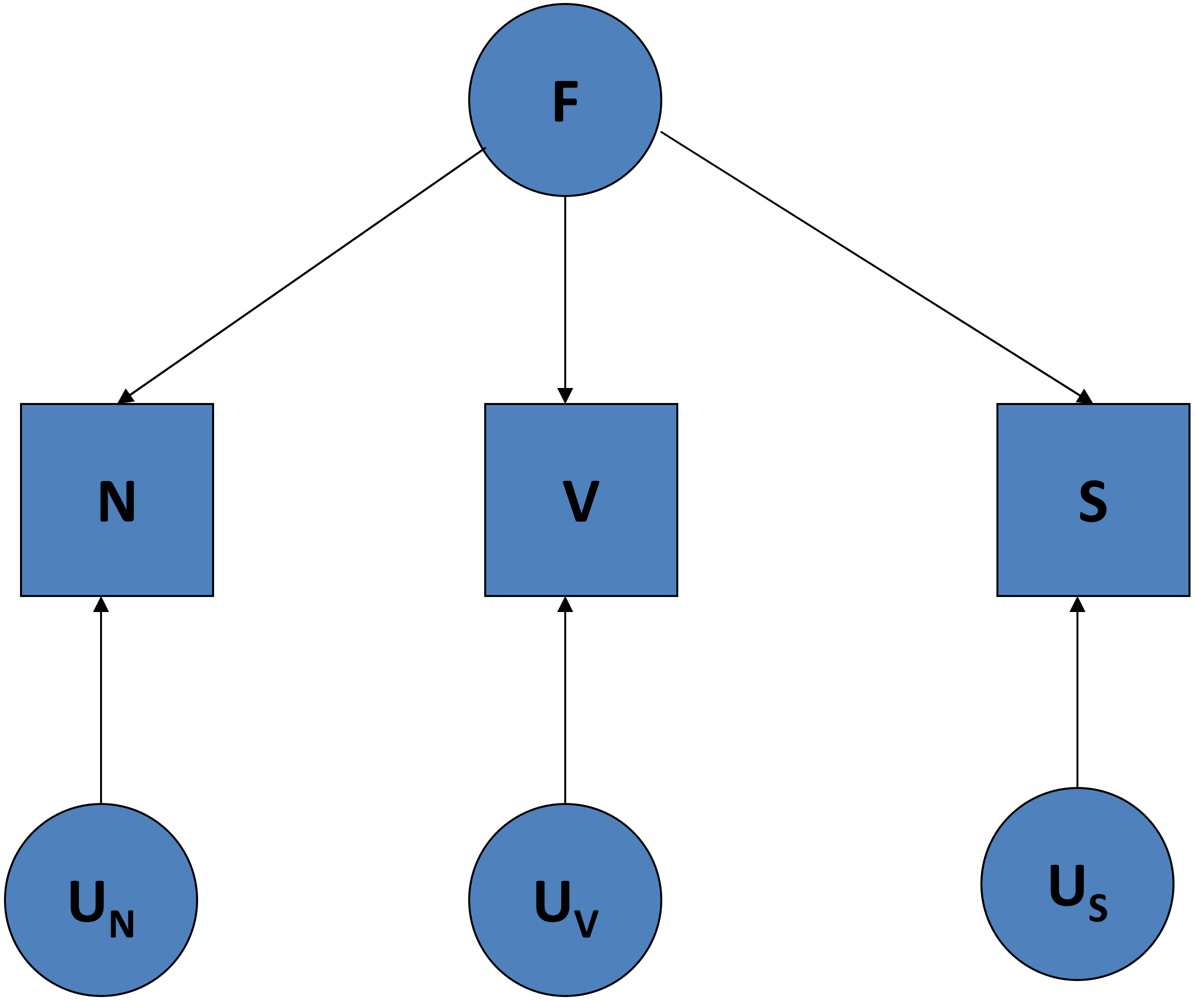
Path diagram for the effects of genes and environment on phenotype

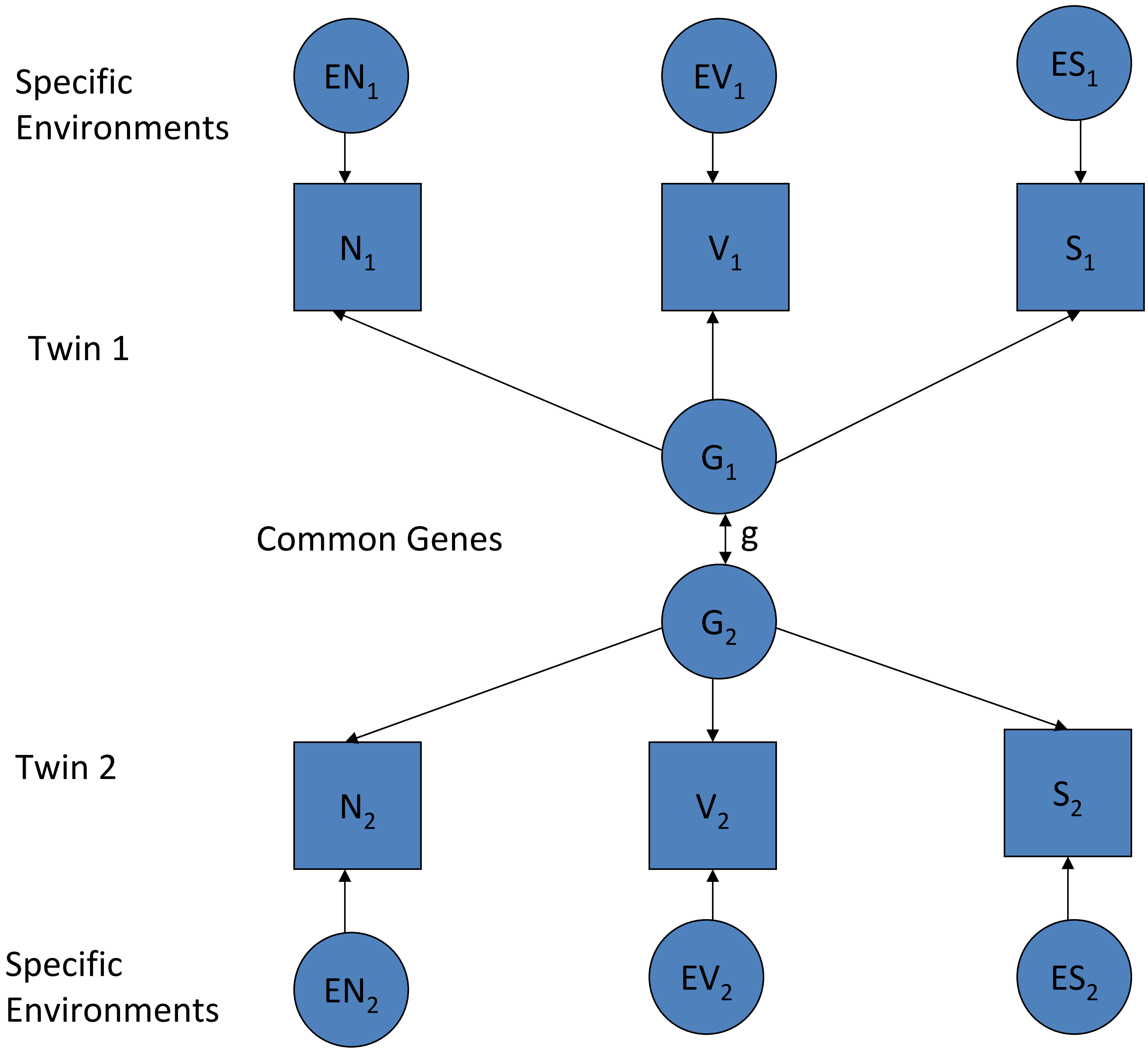


Genetic AND Cultural inheritance?



Multiple Variables

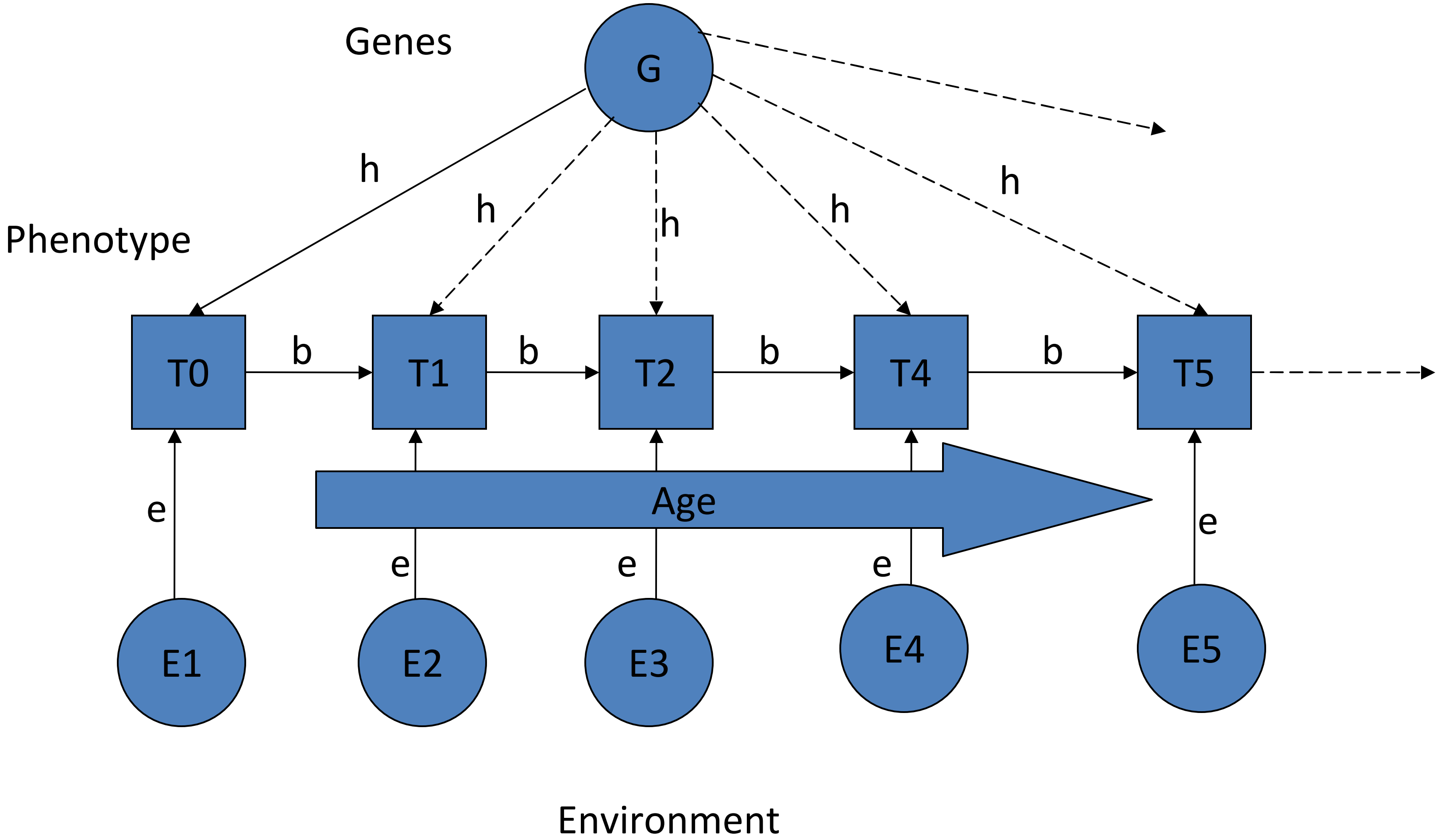

F
N
V
S
U N
U V
U S



Specific Environments
EN_1
EV_1
ES_1
N_1
V_1
S_1
Twin 1
G_1
Common Genes
g
G_2
Twin 2
N_2
V_2
S_2
Specific Environments
EN_2
EV_2
ES_2

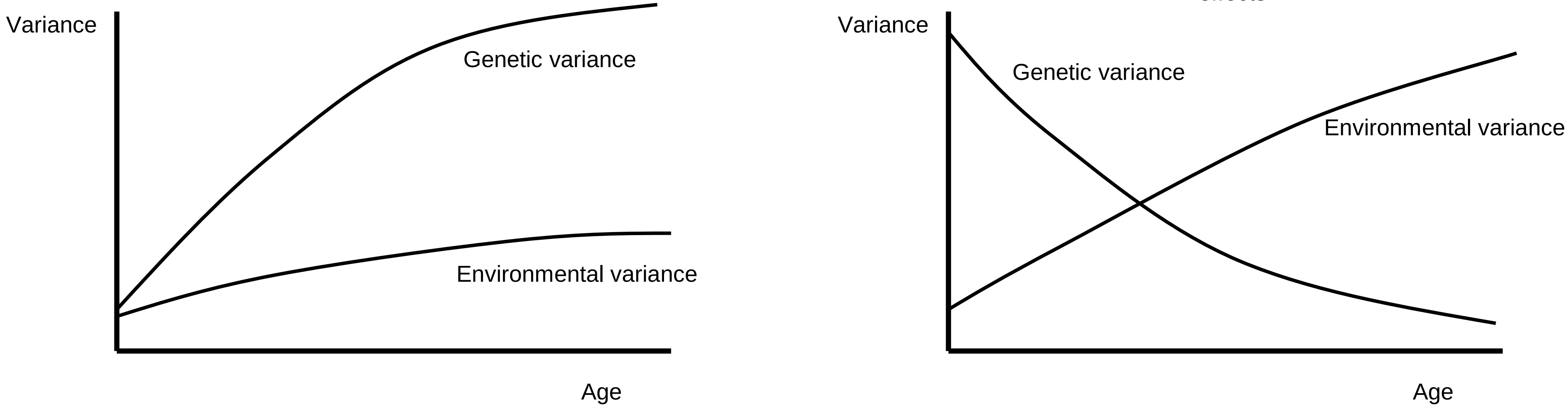
Development

a. Genetic variation in developmental change: time series with common genes and time-specific environmental "innovations"

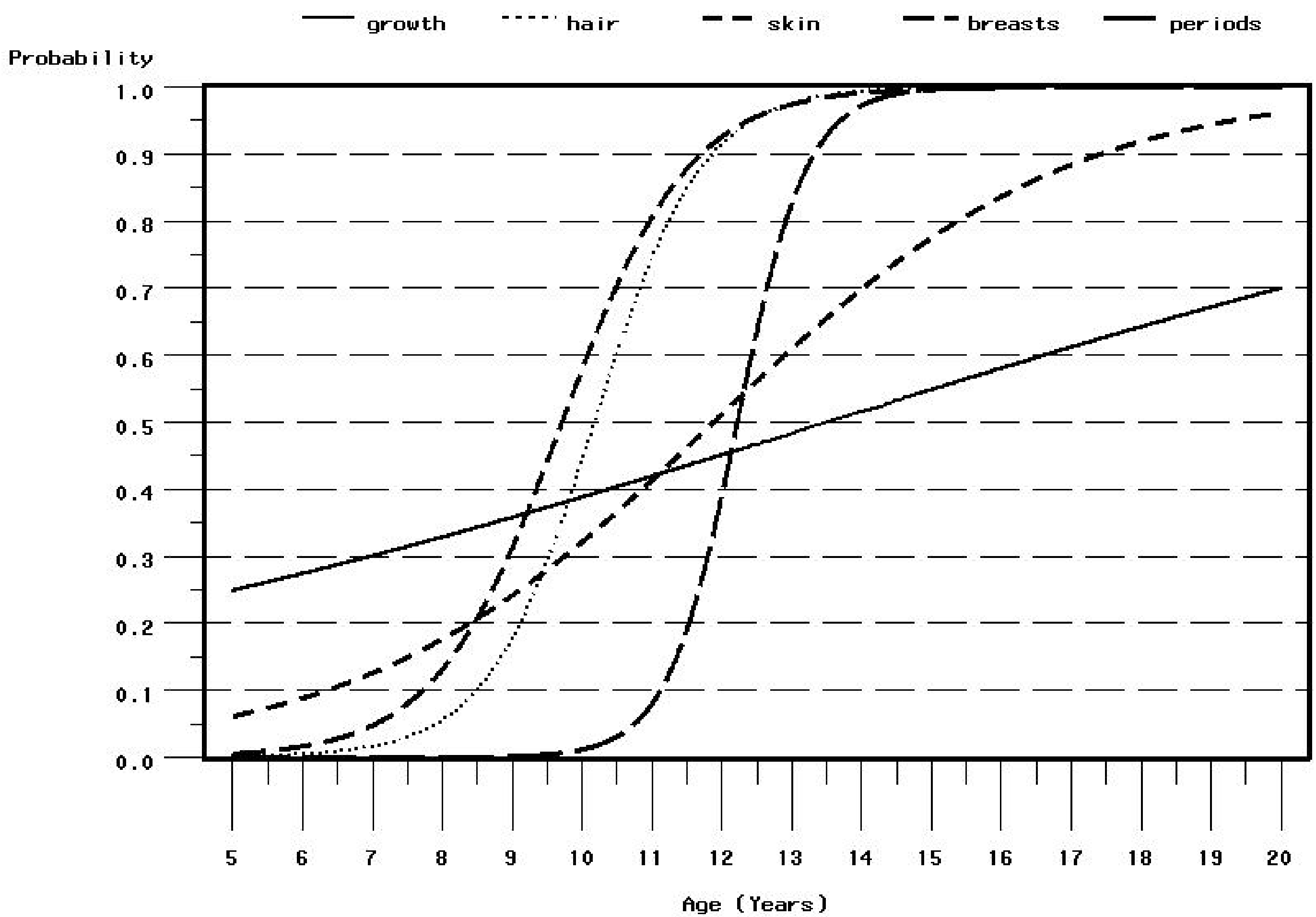


a. Age change in genetic and environmental variance: genetic effect continuous across ages with age-specific environmental effects

b. Age change in genetic and environmental variance: initial genetic effect decays with age with accumulating age-specific environmental effects

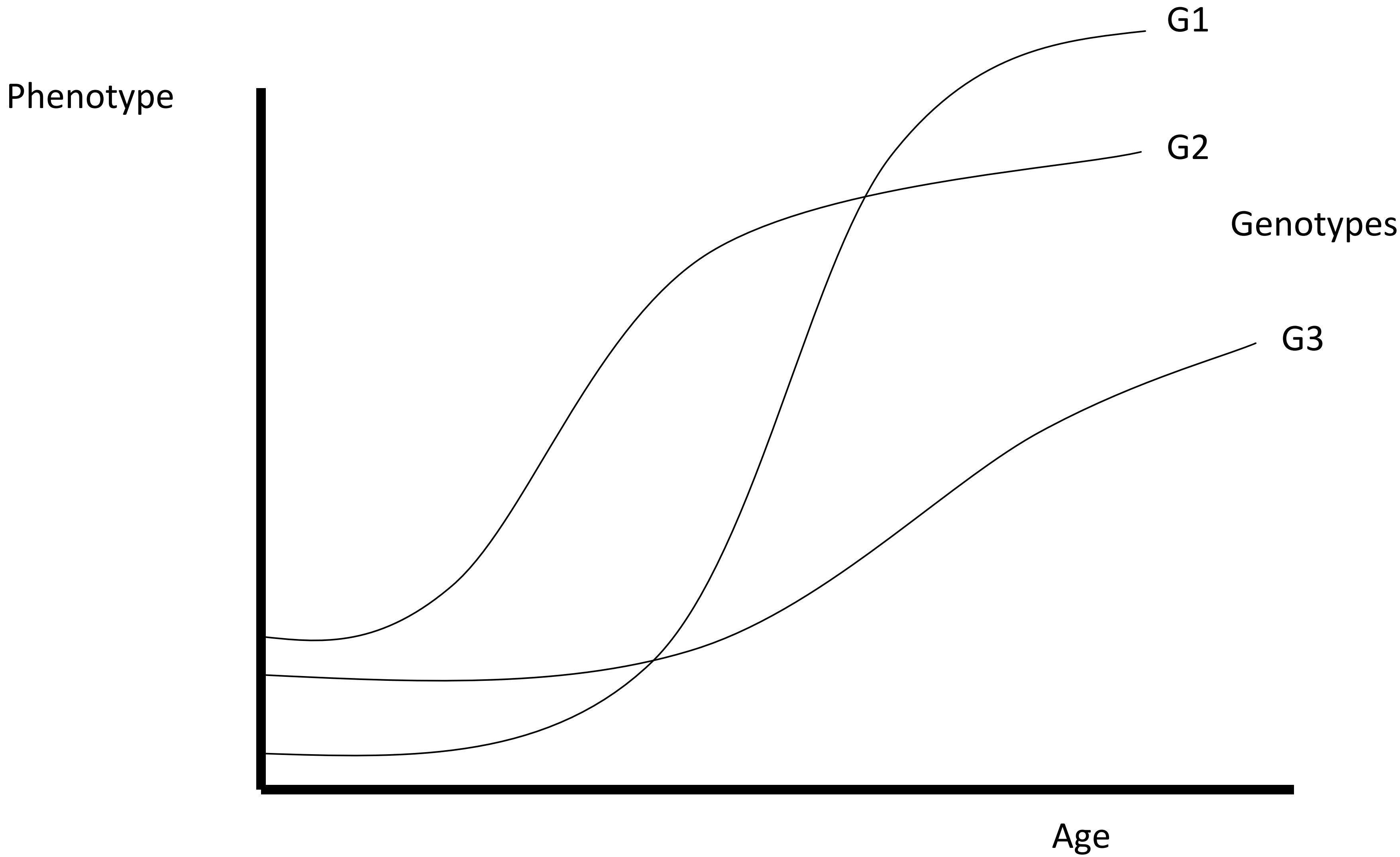


Probabilities of endorsing five puberty-related items in girls .



VTSABD (Eaves et al., 2002)

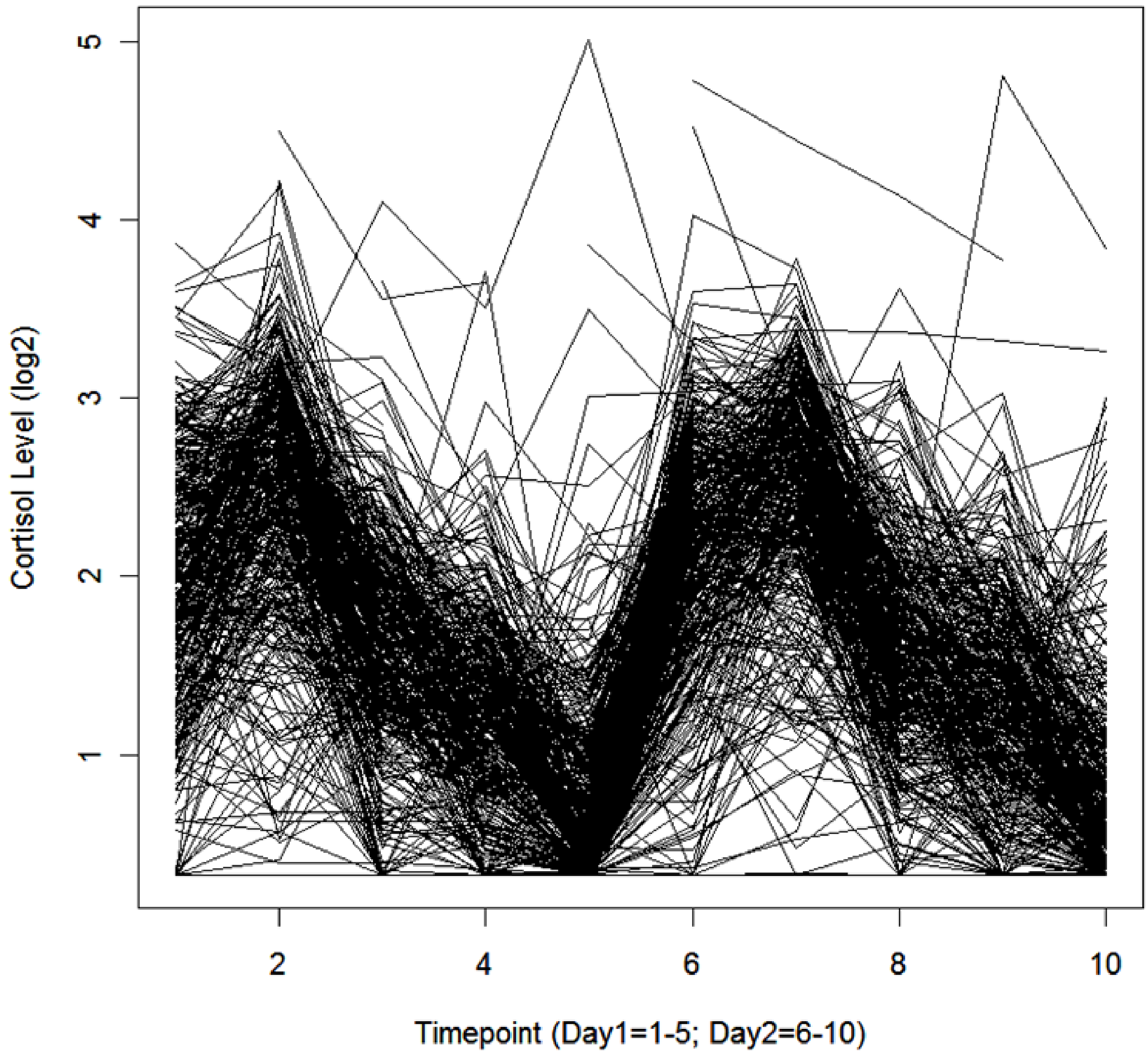
Genetic differences in growth



Random genetic and environmental effects in latent distribution of pubertal advancement/delay relative to age in MZ and DZ girls.

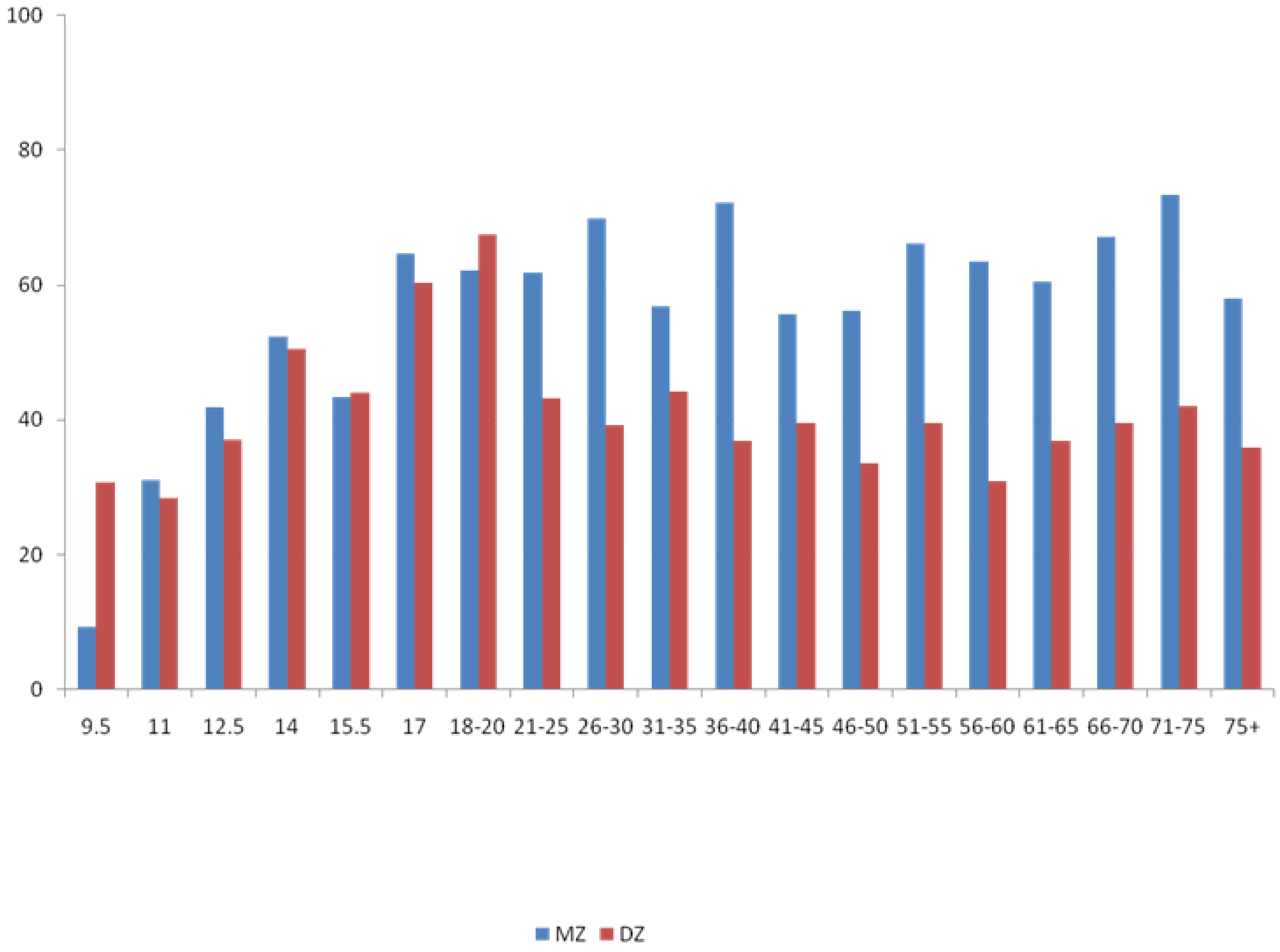
Statistic	Mean	s.d.	2.5%-ile	Median	97.5%-ile
Total Variance	1.171	0.130	0.953	1.155	1.470
% Additive Genetic	76.0	6.4	62.0	76.2	87.5
% Shared Environment	15.4	6.0	5.2	15.4	27.2
% Non-shared environment	8.6	3.1	4.3	8.1	17.7
MZ correlation	0.914	0.031	0.823	0.919	0.957
DZ correlation	0.534	0.035	0.472	0.535	0.603
-2 ln(l)	18880.0	58.52	18770.0	18890.0	19000.0

Data trend for days 1 and 2

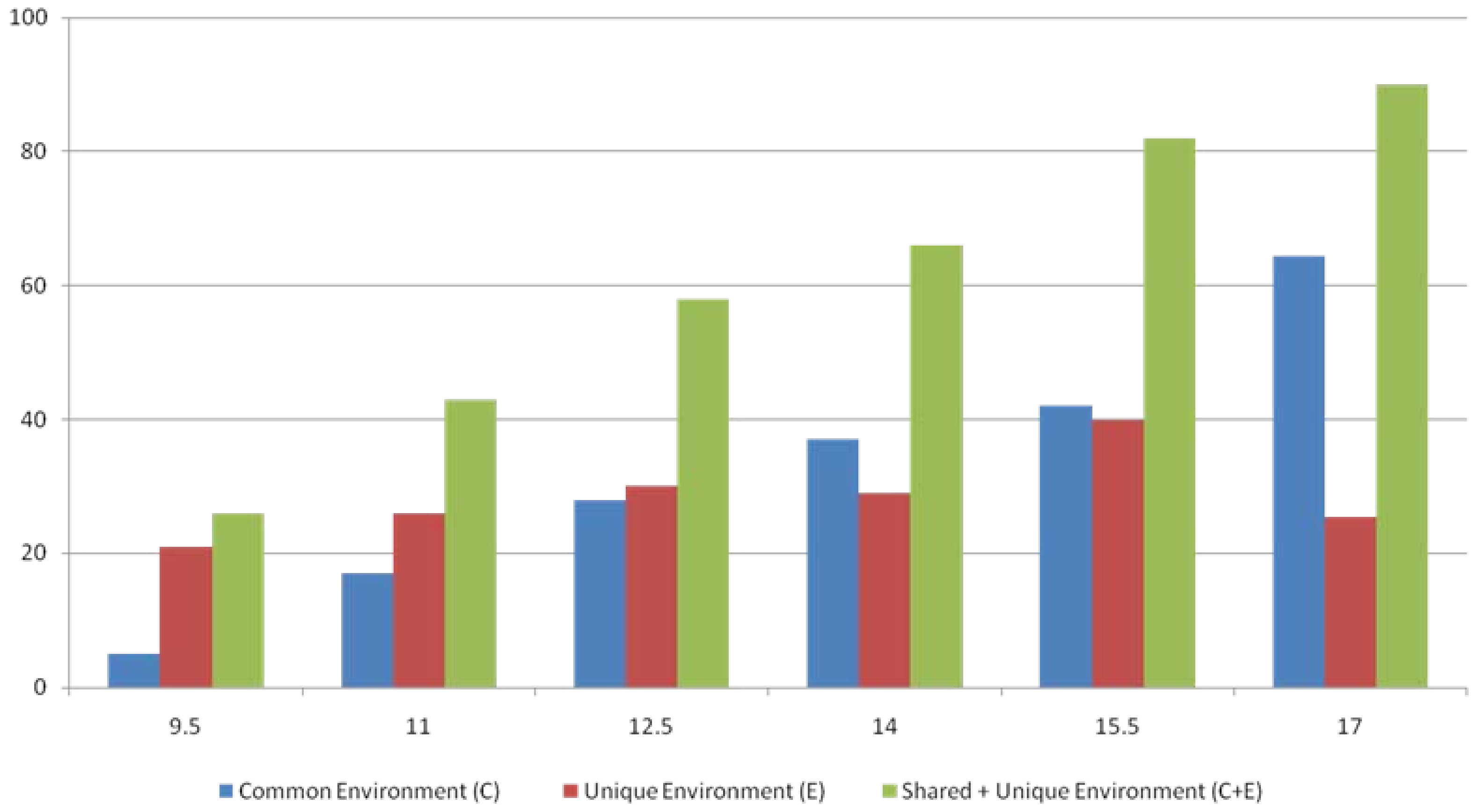


Tim York

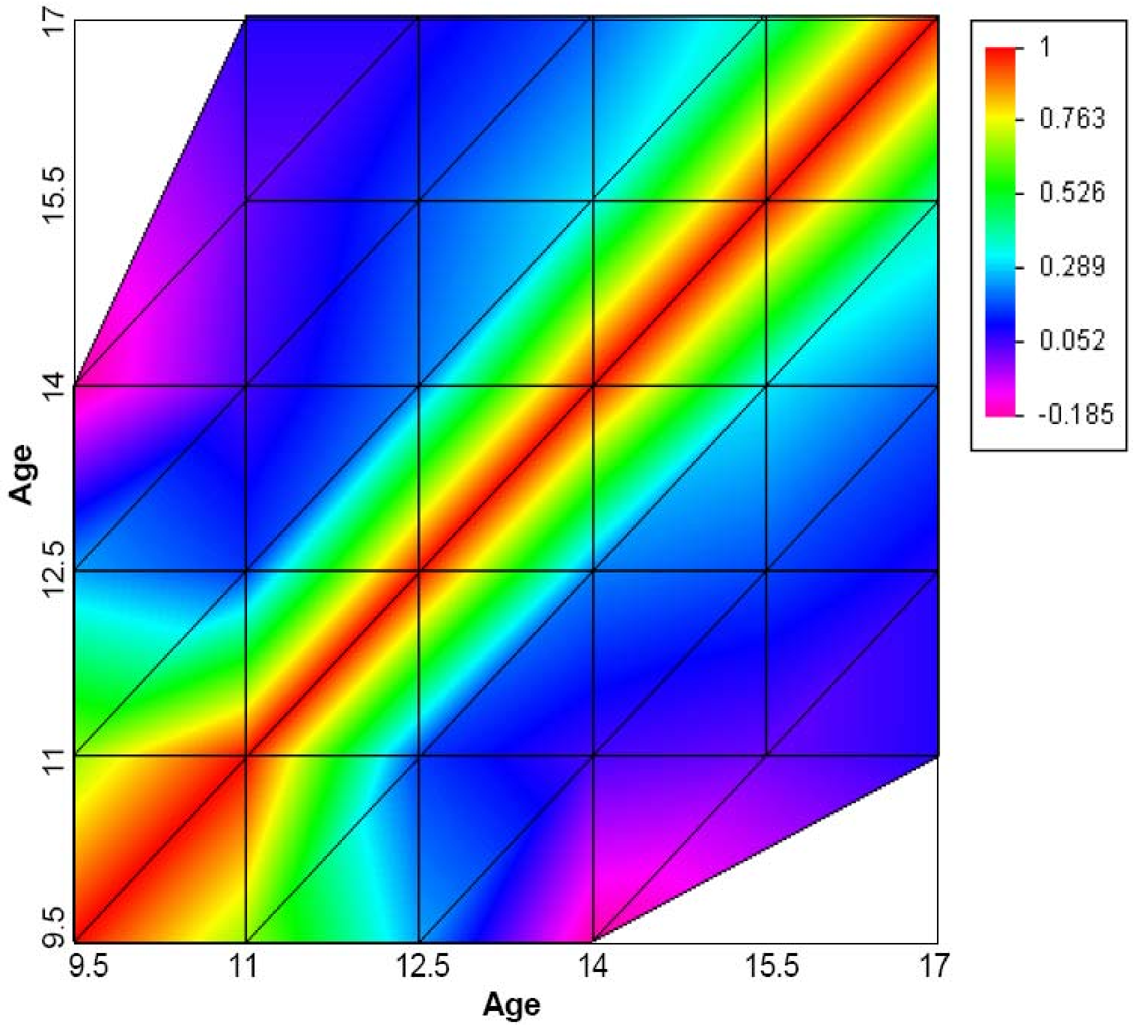
Attitudes over the life-span



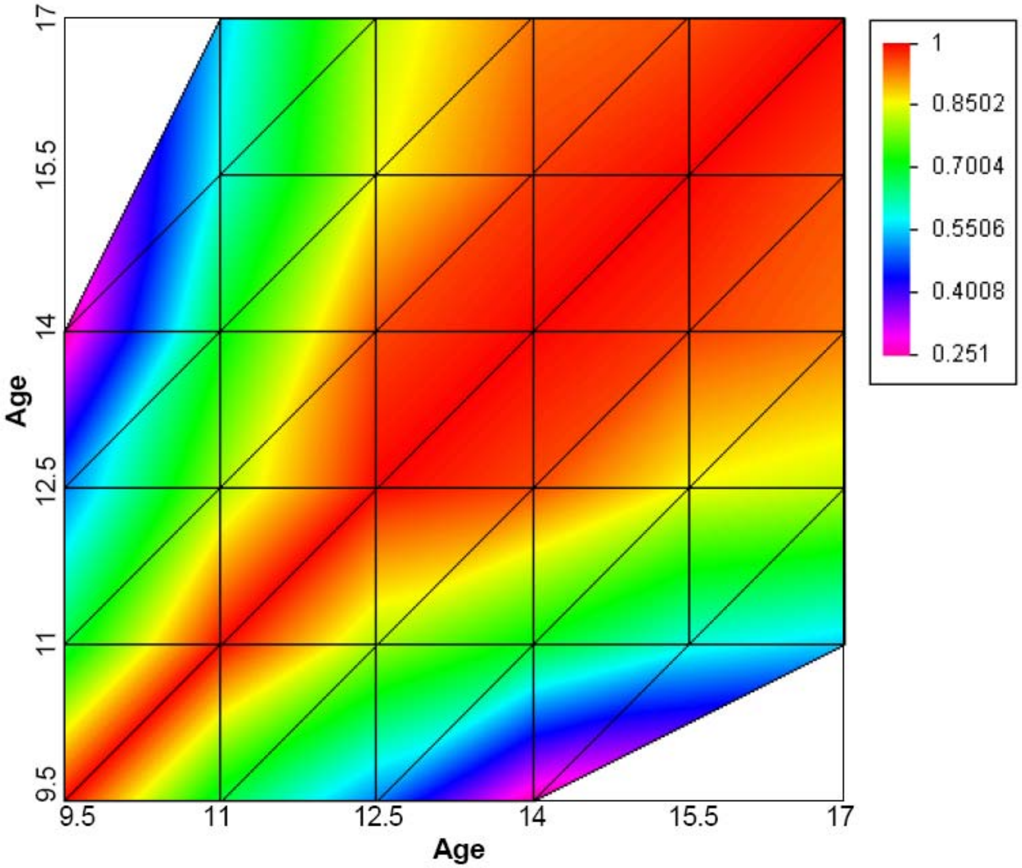
Components of Variance in Adolescence



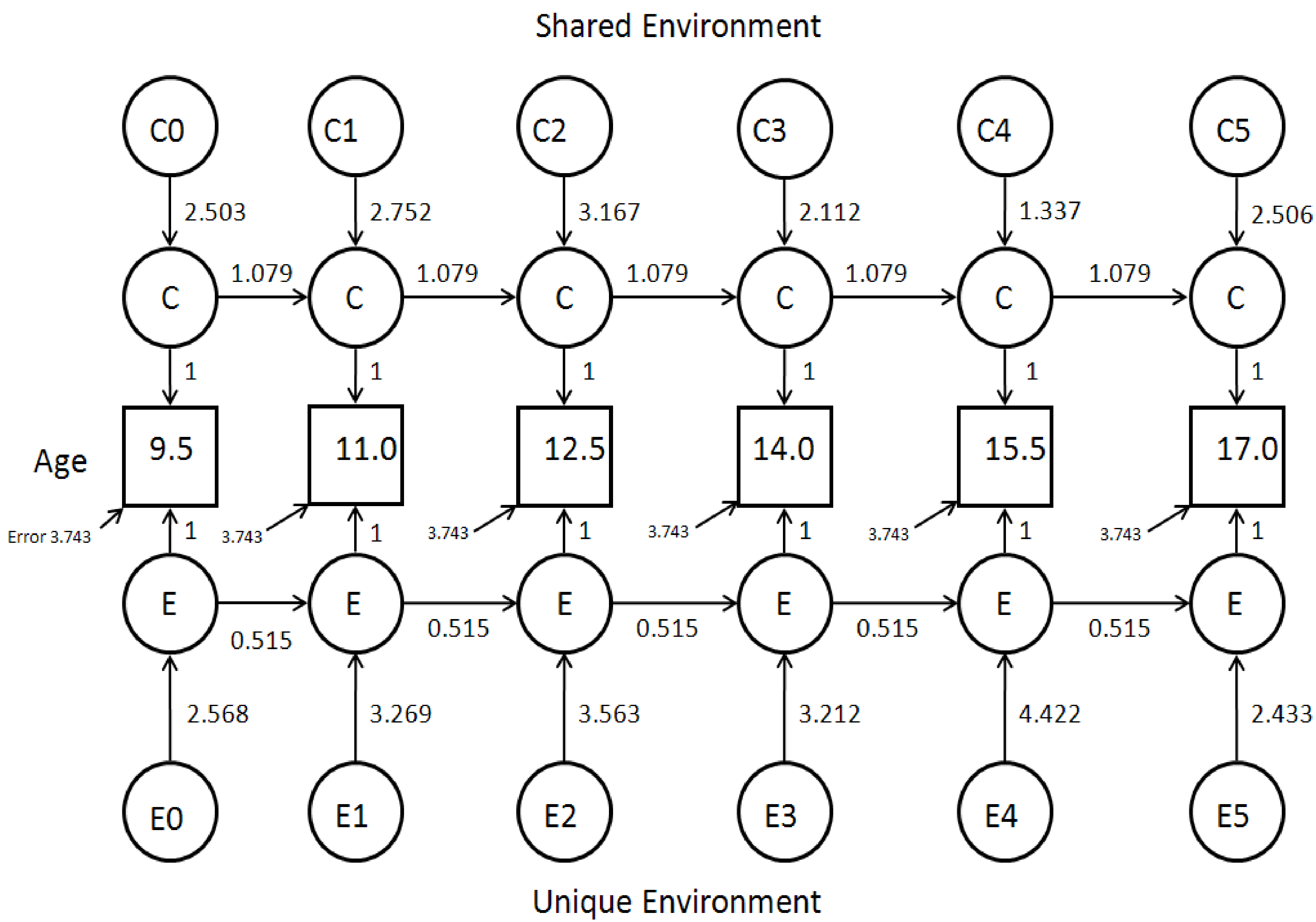
Cross-age correlations in unique environment



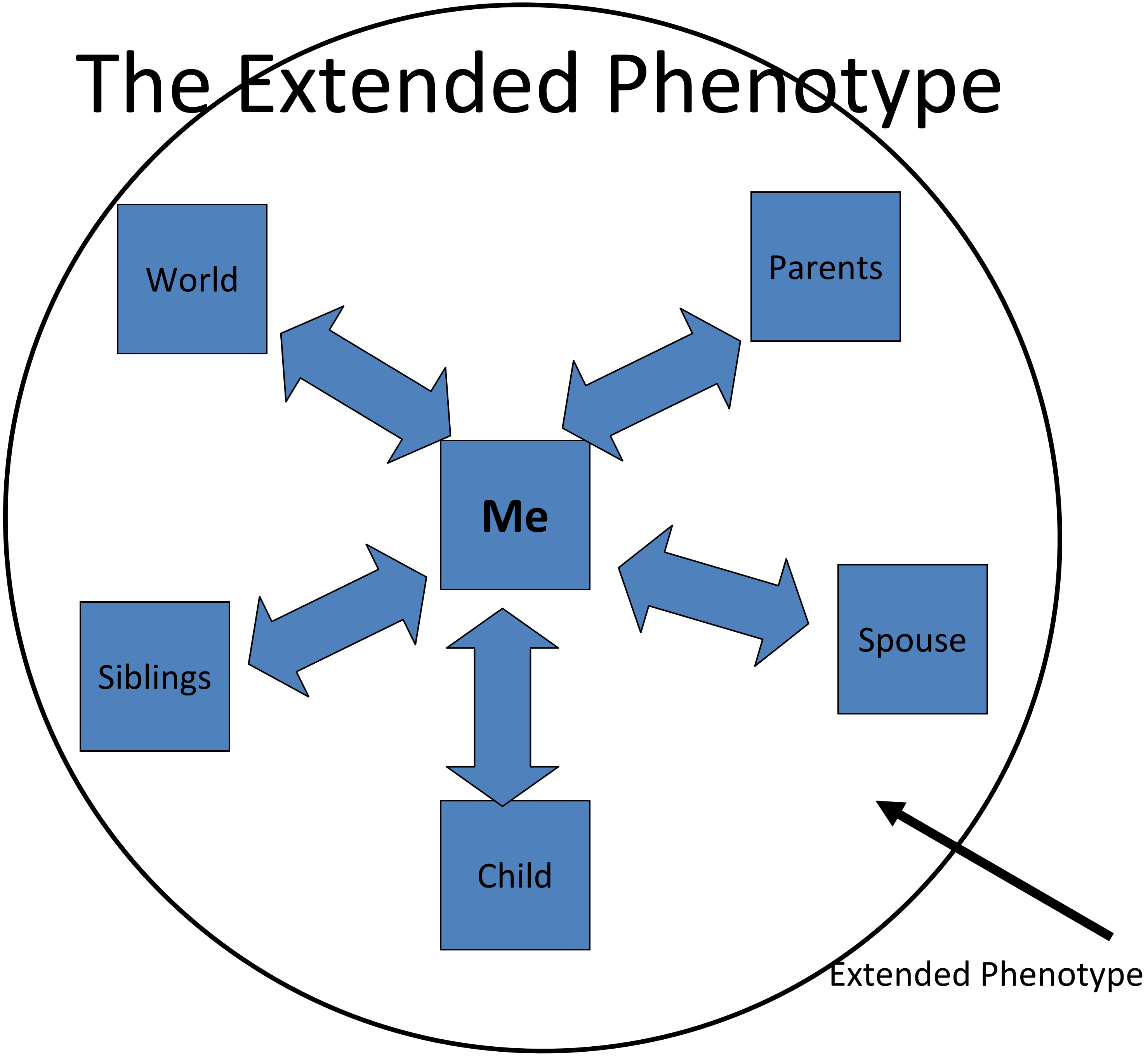
Cross-age correlations in shared environment



Model for Attitude Development

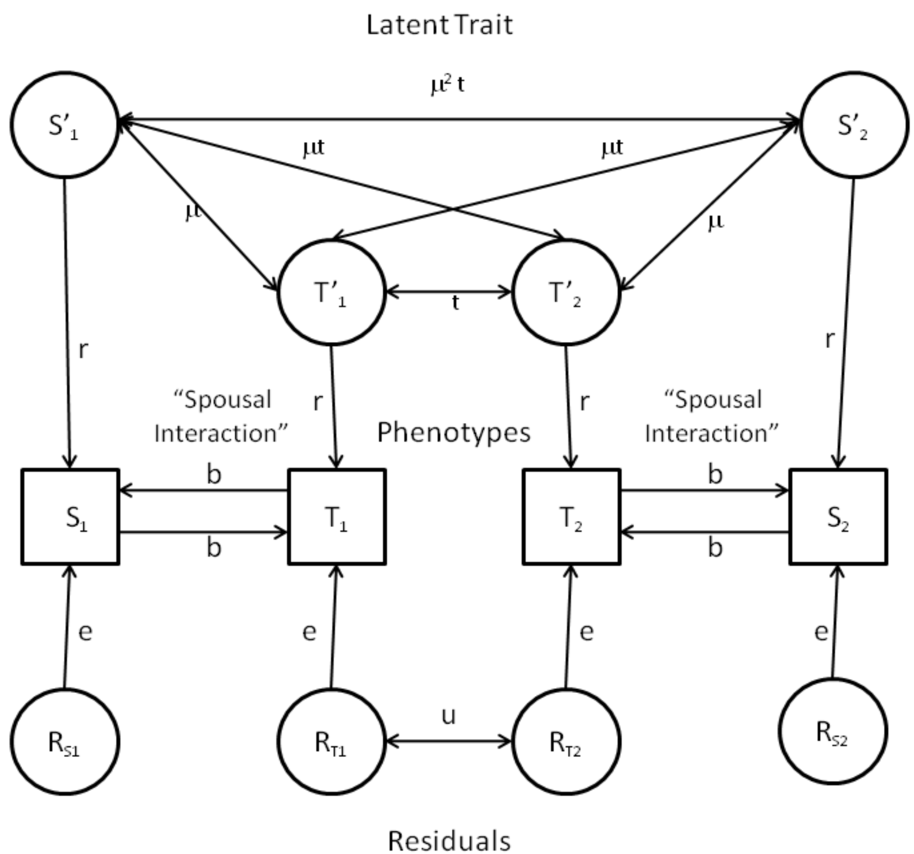


The Extended Phenotype



“Mating”

“Twins and Spouses”



Goodness-of-fit statistics for models for assortative mating in the US and Australia

Model		Random mating	Phenotypic assortment (P)	P+Error	Spousal Interaction	Social Homogamy
d.f.		16	15	13	14	11
Variable	Sample	S^2	S^2	S^2	S^2	S^2
Stature	US	449.179	31.363	24.423[1]	78.930	28.786
	AU	239.827	12.947	11.817[1]	31.694	25.353
Conservatism	US	2535.373	14.845	12.143	118.266	328.491
	AU	2041.407	31.627	29.669	113.276	239.123
Neuroticism	US	63.371	17.811	See note[2]	20.226	19.458
	AU	28.337	17.444	See note[2]	15.583	22.807
Church attendance	US	3375.872	15.187	12.841	103.042	611.006
	AU	3019.544	22.140	21.548[1]	76.574	403.950
Political affiliation	US	2213.625	22.254	18.500	87.889	429.819
	AU	2337.500	34.183	32.537	70.696	322.685
Educational attainment	US	2477.957	46.210	28.207	243.100	57.774
	AU	1430.440	44.146	18.624	160.747	82.086

f(G,E)

Genotype x Environment Interaction

("GxE")

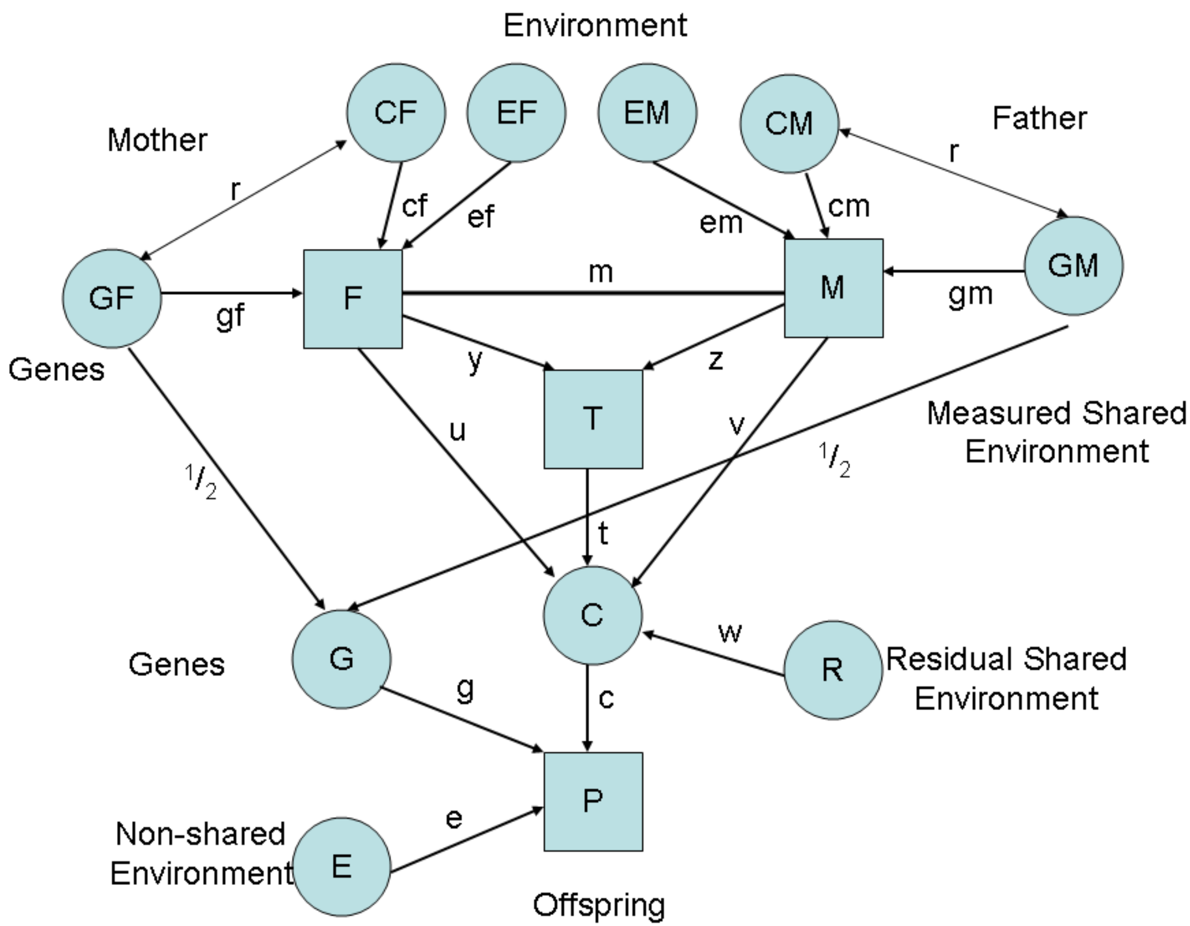
Genotype-Environment Correlation

("rGE")

(“Passive”) rGE

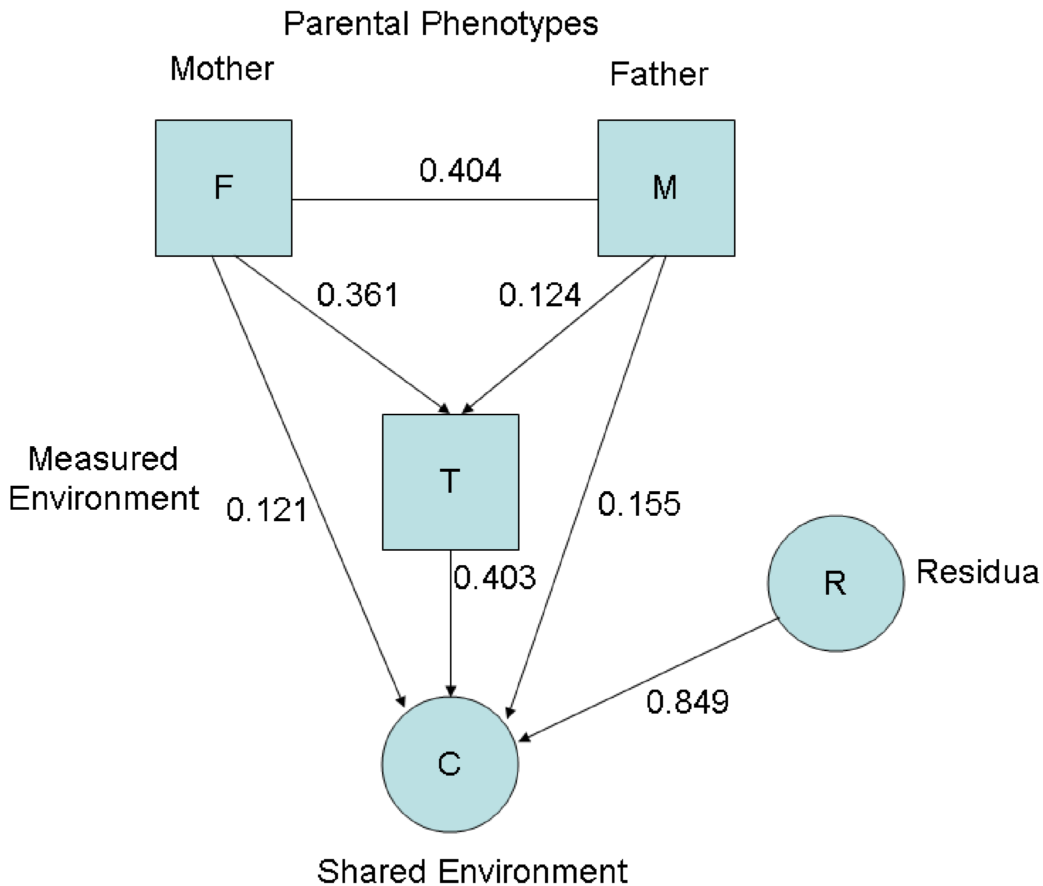
Twins and Parents

Parental Neglect and Anti-Social Behavior

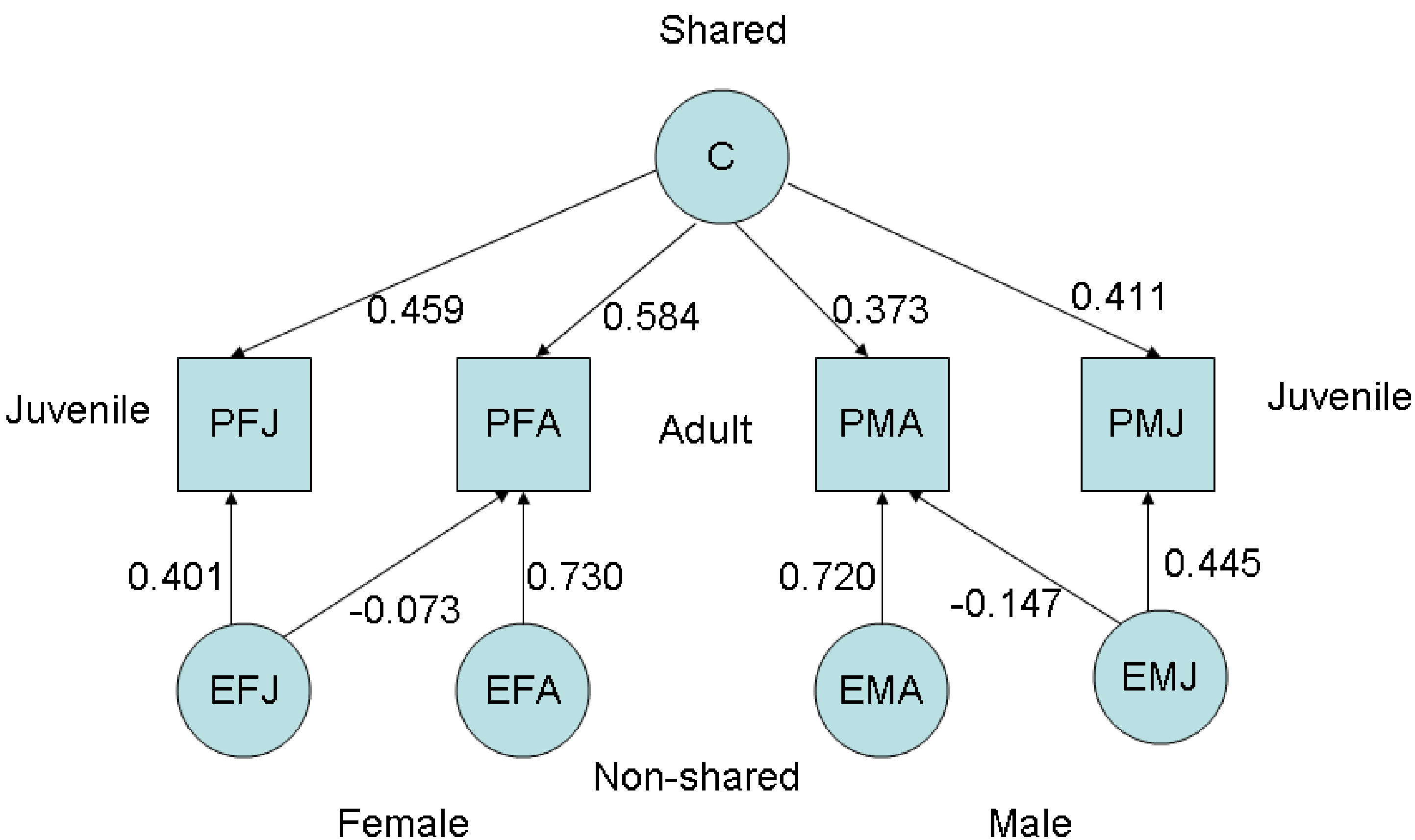


Eaves et al., 2010

Environmental pathways

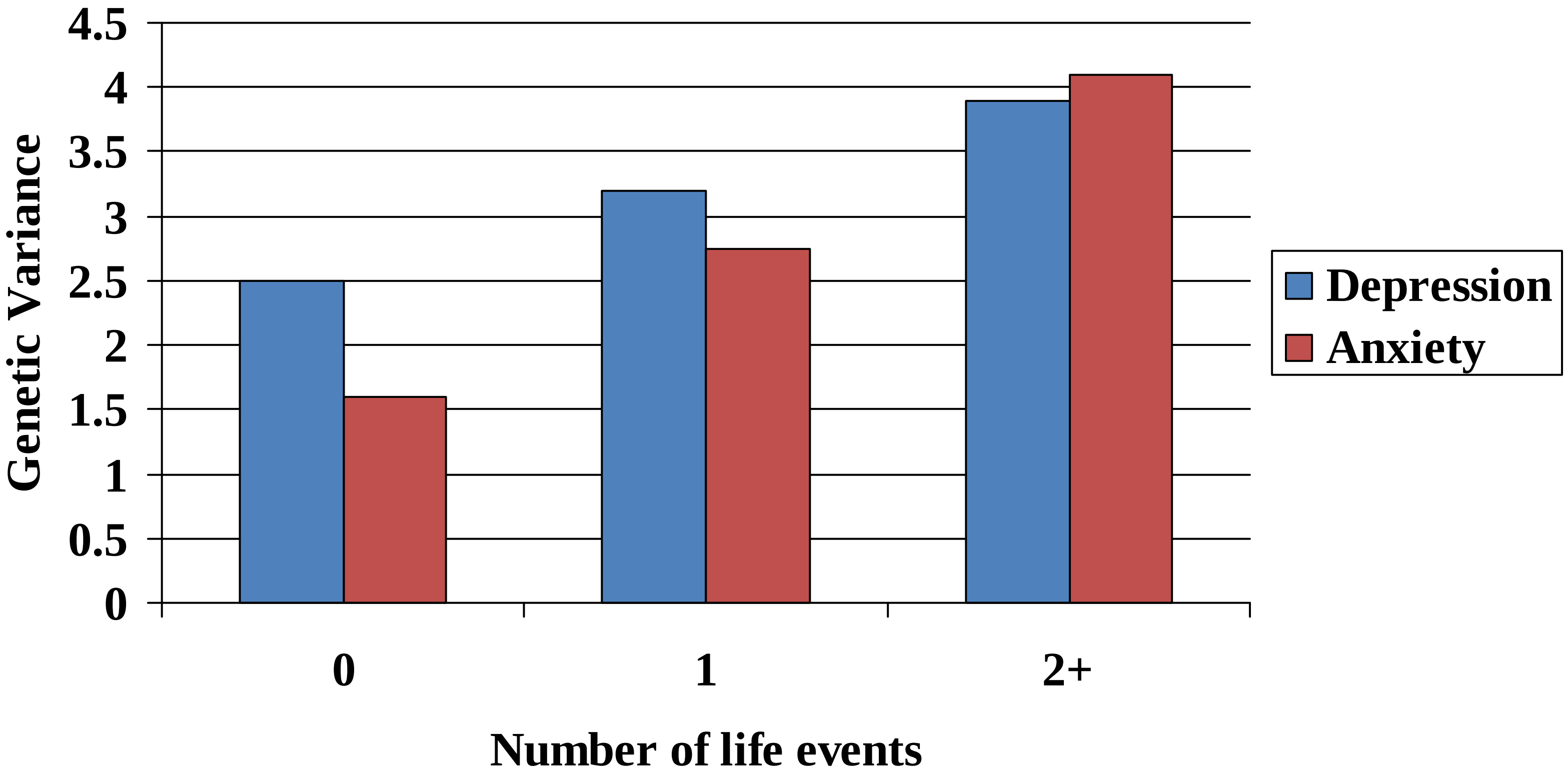


Shared and Unique Environment



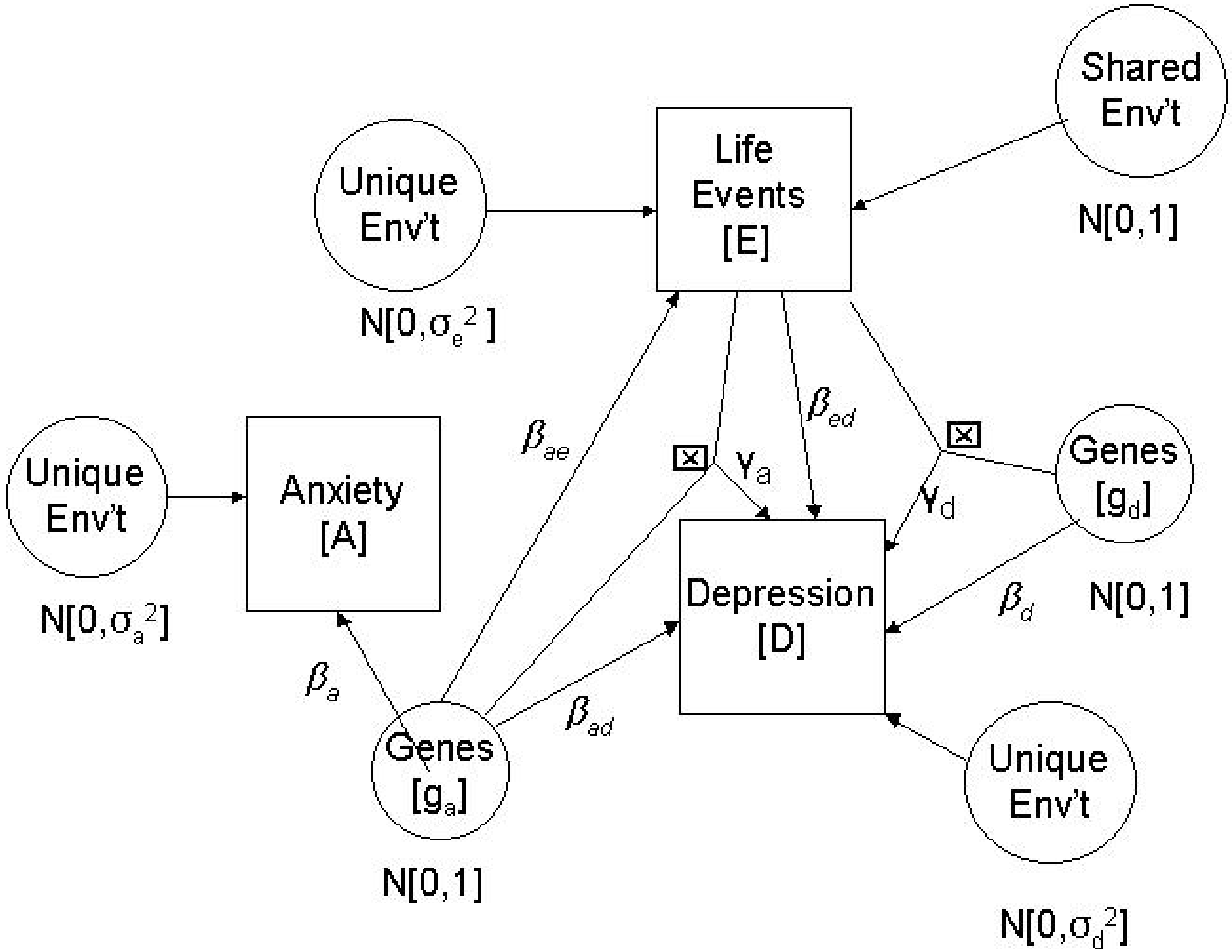
GxE

Genetic Variance and Shared Life Events in Adolescent Females



Putting it all together?

Multiple Genetic Pathways to Depression



Eaves et al. 2003

Have fun!!!