A Polygenic Theory of Schizophrenia

Irving I. Gottesman, James Shields

Proceedings of the National Academy of Sciences of the United States of America, Volume 58, Issue 1 (Jul. 15, 1967), 199-205.

Stable URL:
http://links.jstor.org/sici?sici=0027-8424%2819670715%2958%3A1%3C199%3AAPTOS%3E2.0.CO%3B2-C

A POLYGENIC THEORY OF SCHIZOPHRENIA

By Irving I. Gottesman and James Shields

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, AND
MRC PSYCHIATRIC GENETICS RESEARCH UNIT, LONDON, ENGLAND

Communicated by Theodosius Dobzhansky and read before the Academy April 26, 1967

The mystery of the etiology of schizophrenia has been refractory to the analytical methods provided thus far by the biological and social sciences. Some 50 per cent of the beds in public mental hospitals in the USA are occupied by patients diagnosed as schizophrenics—more than a quarter of a million individuals. Empirical drug treatments now bring symptomatic relief to many of these individuals but no rational therapy founded on etiological knowledge has yet appeared with preventative or curative powers. Strong views exist about the significance of the role to be accorded genetic factors in the etiology of the disorder, ranging from minimal to sufficient. So long as we limit ourselves to a simple Mendelian framework and construe schizophrenia as a homogeneous disease entity, we do not have the "workable concept of heredity" called for by Bleuler[1] in 1911 when he coined the term; it has been known all along that the classical ratios are not found among the relatives of schizophrenic probands.[2] Arguing by analogy with advances in mental deficiency, some subtypes, individually of rare occurrence, might fit simple models; multiple use of the latter strategy would still leave most schizophrenias unexplained, paralleling the mental retardation situation, and so far no Mendelian forms have been identified. At the risk of being labeled Procrustean, ideas about the genotypic homogeneity of the schizophrenias could be preserved by the reasonable use of the concept of incomplete manifestation.[3] To invoke continued mutation as the cause of a disorder with a lifetime morbid risk of eight per thousand puts a strain on current beliefs about the mutation rate for human loci of about 10^{-5}. Evidence for the popular view that environmental hardships are sufficient to account for the etiology of schizophrenia stems from the observation[4] that the prevalence of the condition is eight times higher in the lower social classes than in the upper. The direction of causality implied by the latter finding has been undermined by data in the United Kingdom[5] and the USA[6] showing that when schizophrenics are classified according to their fathers' social statuses, all classes are represented proportionately, thus suggesting that premorbid symptoms of schizophrenia lead to downward social drift. Further evidence against the crucial etiological importance of a poor environment comes from the fact that only some 35 per cent of the offspring of two schizophrenic parents become schizophrenic,[7] and Heston's[8] report of a risk for schizophrenia of 16.6 per cent in children separated from their schizophrenic mothers within three days of birth. The data above together with the frequently documented excess risk for schizophrenia in the twins[9] and other close relatives[10] of index cases make genetic arguments seem most plausible in accounting for the observations, but the nature of such arguments must require some form other than the traditional ones.

One of the possibilities not given sufficient attention involves positing a large proportion of cases as being polygenically determined. We should like to consider the merits of treating schizophrenia as a threshold character whose appearance is

predictable from a diathesis-stress model.[11] Grüneberg's[12] concept of quasi-continuous variation and Lerner's[13] concept of phenodeviants are relevant to our model building. Let us suppose that the diathesis is polygenically determined and that what is inherited is a constitutional predisposition to developing schizophrenia. Descriptions of the diathesis on more than one level, e.g., biochemical, cell membrane, or neuronal fine structure, may be essential in formulating research strategies. While some of the genetic influences may exert themselves early in development, others may be augmented or released only after specific psychosomatic states have been reached.

Polygenic theory would predict the continual appearance of segregants in the offspring of normal parents, increased risks of schizophrenia in highly loaded families, and a slow response to the negative selection associated with lowered marriage and fertility rates.[14] The latter would obviate the need to find the heterozygote advantages implied by a balanced polymorphism view.[15] Polygenic theory could explain the relationship found by ourselves[16] and others[17, 18] between severity of the disorder in twin probands and levels of concordance in their co-twins. We could infer that mild cases with good outcomes had had few of the genes in the system, and we would expect their co-twins to have a lower probability of encountering stress sufficient to cause decompensation than the co-twins of severely ill cases. Irrational and schizoid relatives of probands would be viewed as being near but below the threshold. The latter together with a variety of balancing forces[19] would help maintain the polygenes in the gene pool. Schizophrenics could be thought of as part of the genetic load, the price paid for conserving genetic diversity.[20]

Although schizophrenia is necessarily viewed as an all-or-none character for record-keeping purposes, clinical contact with preschizophrenics or "recovered" cases shows clearly the artificiality of such a dichotomy. Until recently, data in the form of incidences could not be brought into the fold of quantitative genetics,[21] further adding to the inertia of exploring schizophrenia as a polygenic trait, and frustrating attempts to estimate the relative importance of hereditary and environmental factors in its etiology. Falconer[22] (cf. ref. 23) has now made available a model and methods for the handling of disease incidences in the relatives of probands with threshold characters, i.e., diseases that appear to have an all-or-none manifestation but are in fact determined by an underlying graduation of some attribute *really* causing the disease. The latter attribute has been termed *liability* by Falconer and is intended to convey not only the individual's innate tendency to develop the disease (i.e., susceptibility), but also the environmental milieu to which he is exposed that makes him likely to develop the disease. The point on the scale of liability above which all persons are overtly affected is called the threshold. It is the heritability of the liability to schizophrenia that is our chief concern.

Heritability (h^2) expresses the degree to which phenotypes shown by parents are genetically transmitted to their children and is usually estimated from the degree of resemblance between relatives measured as a correlation. Essentially Falconer's method converts incidences into regression coefficients which in turn lead to an estimate of the heritability of liability. Heritability is related to the *degree of genetic determination* in the following way: h^2 is the additive genetic variance as a proportion of the population phenotypic variance; degree of genetic determination is the total genetic variance (additive plus nonadditive) as a proportion of the total

variance. The degree of genetic determination will equal h^2 in the absence of variance from dominance or gene interaction, or it may be greater, but it can never be less than h^2.

Imagine two normal distributions on a base representing a scale of liability, one for a reference population and one, displaced to the right, for the relatives of schizophrenics. We must introduce the following definitions following Falconer: G = mean liability of the general population; A = mean liability of schizophrenics in the general population; R = mean liability of relatives of schizophrenics; q = incidence of schizophrenia; x = distance of the threshold from the mean liability (normal curve deviate units); z = height of ordinate at threshold; and a = mean distance of schizophrenics in the general population ($A - G$) from mean liability of general population ($= z/q$). Falconer provided tables of x and a which are entered with values of q, the incidence of schizophrenia in any degree of relative and the incidence in the general population. For the latter we have used the figure of 1 per cent but have also looked at the results if q should be 2 per cent. An important feature of the method is that it permits the calculation of the standard error of a regression coefficient which can then be converted into the standard error of h^2, a statistic that behavior genetics has lacked until now.

The regression, b, of relatives on probands, is given by

$$b = \frac{R - G}{A - G}, \tag{1}$$

where the denominator is analogous to the selection *differential* and the numerator to the *response* in a selection-type of experiment (ref. 24, p. 65). Equation (1) expressed in terms of normal curve statistics becomes

$$b = \frac{X_{\text{general pop.}} - X_{\text{relatives}}}{a}. \tag{2}$$

The h^2 is readily derived from b as follows: Let P represent the liability of any person, R that of a proband's relative, and r the genetic coefficient of relationship. The regression of R on P is equal to the covariance RP divided by the total population phenotypic variance or r multiplied by the additive variance divided by the phenotypic variance. Since h^2 was defined as the ratio of the latter two variances, we have $b = rh^2$, whence

$$h^2 = b/r. \tag{3}$$

It will be recalled from the laws of Mendelian segregation that $r = 1$ for MZ twins, $1/2$ for DZ twins, sibs, and parent $\times$ child, and $1/4$ for second degree relatives. Falconer has applied the technique to data from relatives other than twins for renal stone disease, congenital pyloric stenosis, club foot, and peptic ulcer, with h^2 values ranging from 37 $\pm$ 6 per cent for ulcer to 79 $\pm$ 5 per cent for pyloric stenosis. Application of the method to schizophrenia gives h^2 values with their standard errors for $q = 1$ per cent and $q = 2$ per cent as shown in Table 1. Two per cent is almost certainly too high for the incidence of schizophrenia generally, but some estimates based on typical rural populations in Norway[25] and Sweden[26] are nearer 2 than 1 per cent. A lifetime incidence of 2 per cent would also make provision for the possibility that for every hospitalized psychotic there is at least one in the general

TABLE 1

HERITABILITY OF THE LIABILITY TO SCHIZOPHRENIA

	Investigator	Incidence		Falconer's h^2 $q = 1\%$	Falconer's h^2 $q = 2\%$
Same-sex twins	Slater (1953)[35]				
	MZ co-twins	28/41	68%	105% ± 8%	104% ± 8%
	DZ co-twins	11/61	18%	106% ± 14%	94% ± 15%
	Gottesman and Shields (1966)[16]				
	MZ co-twins	14/28	50%	87% ± 9%	85% ± 3%
	DZ co-twins	4/34	12%	86% ± 21%	72% ± 23%
	Kringlen (1966)[36]				
	MZ co-twins	28/64	44%	82% ± 6%	79% ± 7%
	DZ co-twins	12/100	12%	86% ± 12%	73% ± 13%
Parents, sibs, and children	Odegaard (1963)[37]				
	Age corrected	84/832	10%	79% ± 4%	64% ± 5%
Sibs	Erlenmeyer-Kimling *et al.* (1966)[38]				
	Observed	131/2007	6.5%	61% ± 3%	45% ± 4%
	Age corrected	131/1260.5	10%	80% ± 2%	66% ± 3%
Aunts and uncles (second-degree relatives)	Odegaard (1963)[37]				
	Age corrected	81/1749	4.6%	96% ± 8%	61% ± 9%

population who, while never hospitalized, was nonetheless psychotic. Such were the findings in the careful epidemiological work of Strömgren[27] on Bornholm (Denmark). The method requires that all independently ascertained index cases be counted and we have done this. It is immediately apparent that the heritability of the liability to schizophrenia, however defined in the various studies, is quite substantial whether q is 1 or 2 per cent and remarkably consistent when estimated from MZ twins, or DZ twins, or siblings, or aunts and uncles. The standard errors, except for DZ twins, are reasonable enough to give confidence to the h^2 values obtained. It is necessary to caution that the heritability is a property not only of the trait but also of the population sampled and its effective environmental milieu. The stability of the values in Table 1 is impressive to us considering that they come from three different cultures, from mild and severe schizophrenics, and from relatives sharing a wide range of communality of environment with the probands.

We readily admit that our findings are far from monolithic. Values of h^2 greater than 100 per cent are obvious signs that the method is subject to errors, or that underlying assumptions have not been fully met. The sharing of trait-relevant environmental factors is a possible source of error that inflates the values. This error is especially likely in siblings, whether they are twins or not. Another source of error in the estimates of h^2 arises only with DZ twins and ordinary sibs. Doubling this particular regression coefficient leads to an estimate of the additive genetic variance, together with $1/2$ of the nonadditive variance arising from dominance, as a proportion of the total variance. Thus, the presence of appreciable nonadditive variance leads to an exaggerated value of h^2 when the regression coefficient is doubled, but the value obtained is still less than the degree of genetic determination. An evaluation of the relative importance of nonadditive genetic variance can be obtained, however, by subtracting the regression for DZ twins from that for MZ twins. The remainder estimates $1/2$ of the additive genetic variance plus $3/4$ of the dominance variance. Applying this procedure to our own data and using an independent estimate of h^2 (i.e., additive variance that could then be halved) from Odegaard's[25] second-degree relatives showed that, at this rough stage of model development, nonadditive variance was relatively unimportant for estimates of the heritability of the liability to schizophrenia. The error from sharing trait-relevant environment also

appears to be unimportant as evidenced by the comparability of h^2 estimates obtained from aunts and uncles with those obtained from twins or sibs. (Falconer[28] doubts that his method can be applied directly to MZ twins since the distribution of liability in their co-twins is skewed rather than normal, thus leading to overestimates of heritability. The fair homogeneity of h^2 values from twins in Table 1 with other relatives probably results from a compensating error on our part that leads to underestimates of twin concordance rates, namely the use of data not age-corrected for co-twins still within the period of risk.)

Falconer heritability estimates will be erroneous if the assumption of a continuous distribution of liability is untenable. Such a situation could arise from incidence data gathered from an unrepresentative sample of schizophrenics, overloaded with severe cases. In addition, a dominant gene with incomplete penetrance might cause a discontinuity in the distribution of liability to schizophrenia; its presence would be inferred if the estimated h^2 were "very obviously too high to be credited" (ref. 22, p. 69). If a threshold trait occurs in the general population with a frequency of 0.1 per cent, an incidence in first-degree relatives greater than 8 per cent yields a value of h^2 greater than 100 per cent with the Falconer formulae; if the general population incidence is 1 per cent, incidences in the same relatives greater than 15 per cent has the same effect on h^2 values.

Up to now, estimates of the heritability of apparently discontinuous disorders such as schizophrenia have depended, for want of anything better, on a formula[29] which did not take the general frequency of the condition into account and which simultaneously treated the concordance rates in MZ and DZ twins. The old formula would yield values of 44 and 61 per cent, respectively, for our twin data and those of Slater in Table 1. The present Falconer formula offers an alternative method based, as we have seen, on the assumption of an underlying normally distributed attribute. Estimates can be based on MZ and DZ twin concordances independently, and it is gratifying to have these match. An advantage of the method is that it permits a quantitative evaluation of the general impression one has about the importance of genetic factors when confronted in schizophrenia by such seemingly low incidences as 6.5 per cent in sibs or 50 per cent in MZ co-twins. It takes into account the lifetime expectation of schizophrenia in the general population of around 1 per cent. Even if we were to take as replicable our concordance rate of 17 per cent in the MZ co-twins of *mild* schizophrenics,[16] the Falconer h^2 would be 51 per cent. Our values of the heritability of the "liability" to schizophrenia, would appear to be as high as, if not higher than, those calculated for congenital abnormalities and physical diseases of later onset that are commonly regarded as having a strong hereditary component.

Diabetes mellitus is a common disorder with an obscure etiology but whose manifestations are more palpable and tractable in a laboratory than those of schizophrenia. Diabetes has been termed "a geneticist's nightmare";[30] it shares a number of other parallels with the problems of researching schizophrenia also. The incidences in the general population and among the relatives of index cases are similar; it is not clear whether diabetes is a qualitative departure from normal glucose metabolism or simply the tail of a normal distribution;[31, 32] and the frequency with which the disorder is diagnosed is strongly influenced by trait-relevant environmental variables such as general nutritional level.

The data from diabetic twins are of most interest to us. Then-Bergh,[33] in a

1938 paper, found concordance rates for clinical diabetes in MZ and DZ twins to be 43 and 14 per cent, respectively, in pairs over the age of 43, thus minimizing the need for age correction. If either overt diabetes or abnormal glucose tolerance test results were taken as indicators of the diabetic genotype, the two concordances rose to 100 and 40 per cent. More recently Harvald and Hauge[34] using the Danish Twin Registry found concordance rates for overt diabetes in MZ and DZ twin pairs of 64 and 9 per cent, respectively, where pairs with healthy co-twins under the age of 40 had been omitted (at our request) to make the findings comparable with the data above.[33] Heritability estimates using the incidences in co-twins of index cases range from 58 ± 13 per cent to 100 ± 6 per cent based on a population incidence of 2 per cent. In our opinion some research strategies found useful in diabetes may be profitably adapted and applied to schizophrenia.

Falconer[22] has been careful to point out that high heritabilities mean that the environmental factors operating in the populations sampled are unimportant as causative agents of the disease, but that this provides no information about the potential for curing or preventing the disease. "All that could be said . . . is that one will have to look outside the range of normal environments experienced by the untreated population" (ref. 22, p. 69). A polygenic theory and values of heritability based upon it are not explanations one should be satisfied with for long. The task ahead is to identify some of the specific contributing genetic factors and to explore how they interact with other such factors and the environment. If one espouses a polygenic model, it is in hope that it will prove a springboard for further advances and not a replacement for other viable theories. Nature appears to have retained for itself a flexibility that permits a trait to evolve from monogenic control to polygenic control and vice versa (ref. 19, p. 41). We must learn to tolerate the Protean qualities of schizophrenia, to eschew the Procrustean, and to embrace the Promethean, whatever their source.

Summary.—The etiology of schizophrenia is unknown, but strong views exist about the significance of the role to be accorded genetic factors, ranging from minimal to sufficient. Proponents of a genetic etiology have usually talked in terms of Mendelian genetics. Classical segregation ratios are not found for the relatives of schizophrenics. One of the possibilities not given sufficient attention involves positing a large proportion of cases as being polygenically determined. Thus the disorder would be treated like a threshold character (Grüneberg) whose phenotypic appearance would depend on both the number of genes present and the amount of stress. The model would predict the continual appearance of segregants in the offspring of normal parents (Lerner's phenodeviants), increased risks of schizophrenia in loaded families, and a slow response to negative selection. Schizophrenics could be thought of as part of the genetic load, the price paid for conserving genetic diversity.

Falconer has now provided techniques for estimating the heritability of the liability to human diseases that appear to have an all-or-none manifestation but are in fact determined by an underlying graduation of some attribute really causal. Essentially the method converts incidences into regression coefficients and the latter into estimates of the heritability of liability. Heritability was substantial (greater than 60%) when estimated independently from identical or fraternal twins, siblings, parents, and aunts and uncles. Similar results were obtained for data from the

relatives of diabetics. A valid polygenic model permits the conclusion that relatively low incidences of schizophrenia (e.g., 5%) in the parents or siblings of probands is compatible with a high degree of genetic control.

We wish to thank D. S. Falconer, Edinburgh, for his critical comments on our ideas, and B. Harvald and M. Hauge of the Danish Twin Registry, Copenhagen, for unpublished data on diabetes. We wish to acknowledge the technical assistance of Beatrice Rouse, University of North Carolina, and the encouragement and foresight of E. Slater, Director, Medical Research Council Psychiatric Genetics Research Unit, London, that have made our work possible. The research was supported in part by USPHS grants MH-13117 and T1-MH-10679.

[1] Bleuler, E., *Dementia Praecox or the Group of Schizophrenias* (New York: Internat. Univ. Press, 1950).

[2] Gregory, I., *Am. J. Psychiat.*, **116**, 961 (1960).

[3] Slater, E., *Acta Genet. (Basel)*, **8**, 50 (1958).

[4] Hollingshead, A. B., and F. C. Redlich, *Social Class and Mental Illness* (New York: Wiley, 1958).

[5] Goldberg, E. M., and S. L. Morrison, *Brit. J. Psychiat.*, **109**, 785 (1963).

[6] Dunham, H. W., Patricia Phillips, and Barbara Srinivasan, *Am. Sociol. Rev.*, **31**, 223 (1966).

[7] Rosenthal, D., *J. Psychiat. Res.*, **4**, 169 (1966).

[8] Heston, L. L., *Brit. J. Psychiat.*, **112**, 819 (1966).

[9] Gottesman, I. I., and J. Shields, in *Progress in Experimental Personality Research*, ed. B. A. Maher (New York: Academic Press, 1966), vol. 3, pp. 1–84.

[10] Shields, J., and E. Slater, *Hosp. Med.*, **1**, 579 (1967).

[11] Rosenthal, D., ed., and colleagues, *The Genain Quadruplets* (New York: Basic Books, 1963).

[12] Grüneberg, H., *J. Genet.*, **51**, 95 (1952).

[13] Lerner, I. M., *Genetic Homeostasis* (New York: Wiley, 1954).

[14] Erlenmeyer-Kimling, L., and W. Paradowski, *Am. Naturalist*, **100**, 651 (1966).

[15] Huxley, J., E. Mayr, H. Osmond, and A. Hoffer, *Nature*, **204**, 220 (1964).

[16] Gottesman, I. I., and J. Shields, *Brit. J. Psychiat.*, **112**, 809 (1966).

[17] Kallmann, F. J., *Am. J. Psychiat.*, **103**, 309 (1946).

[18] Inouye, E., *Proceedings of the Third World Congress on Psychiatry* (Montreal: Univ. of Toronto Press, 1963), vol. 1, pp. 524–530.

[19] Lerner, I. M., *The Genetic Basis of Selection* (New York: Wiley, 1958).

[20] Dobzhansky, T., *Proc. Intern. Congr. Genet., 11th* (1964), p. 541.

[21] Falconer, D. S., *An Introduction to Quantitative Genetics* (New York: Ronald, 1960).

[22] Falconer, D. S., *Ann. Human Genet.*, **29**, 51 (1965).

[23] Crittenden, L. B., *Ann. N.Y. Acad. Sci.*, **91**, 769 (1961).

[24] Fuller, J. L., and W. R. Thompson, *Behavior Genetics* (New York: Wiley, 1960).

[25] Odegaard, O., *Psychiat. Quart.*, **20**, 381 (1946).

[26] Larsson, T., and T. Sjögren, *Acta Psychiat. Scand.*, Suppl. 89 (1954).

[27] Strömgren, E., *Congr. Intern. Psychiat., V, Paris* (1950), p. 155.

[28] Falconer, D. S., personal communication (1966).

[29] Neel, J. V., and W. J. Schull, *Human Heredity* (Chicago: Univ. Chicago Press, 1954).

[30] Neel, J. V., S. S. Fajans, J. W. Conn, and Ruth Davidson, in *Genetics and Epidemiology of Chronic Diseases*, USPHS pub. 1163 (1965), p. 105.

[31] Thompson, G. S., *J. Med. Genet.*, **2**, 221 (1965).

[32] Vallance-Owen, J., in *Advances in Metabolic Disorders*, ed. R. Levine and R. Luft (New York: Academic Press, 1964), vol. 1, pp. 191–215.

[33] Then-Bergh, H., *Archiv. Rassen Ges. -Biol.*, **32**, 289 (1938).

[34] Harvald, B., and M. Hauge, personal communication (1966).

[35] Slater, E., *Psychotic and Neurotic Illnesses in Twins*, Medical Research Council Special Report, Ser. no. 278 (London: H. M. S. O., 1953).

[36] Kringlen, E., *Psychiatry*, **29**, 172 (1966).

[37] Odegaard, O., *Acta Psychiat. Scand.*, **39**, Suppl. 169, 94 (1963).

[38] Erlenmeyer-Kimling, L., J. D. Rainer, and F. J. Kallmann, in *Psychopathology of Schizophrenia*, ed. P. H. Hoch and J. Zubin (New York: Grune and Stratton, 1966), pp. 252–276.