THE INHERITANCE OF NEUROTICISM: A REPLY

By

H. J. EYSENCK, B.A., Ph.D.
Institute of Psychiatry, University of London

THIS is a reply to a critique by Karon and Saunders (9) of the Eysenck and Prell work on the inheritance of neuroticism (6). This paper is a curious combination of sophisticated statistical analysis and argument by irrelevant association. Thus the authors mention the fact that we found a strong heredity predisposition but go on to say that "the results are so much at variance with general clinical experience that doubts arise in the minds of many psychologists . . . particularly those who have investigated, in a therapeutic situation, the source and development of neurotic reactions". The authors do not indicate how it is possible through general clinical experience or in the therapeutic situation, to find an answer to one of the most difficult and complex problems in the whole field of psychology. Whether neurotic predisposition is largely inherited or has little hereditary basis, would not seem to be capable of being discerned simply by giving psychotherapy to a few neurotics.*

Again the authors conclude that "the possible implications of the Eysenck-Prell study are of great practical importance to the practising clinician"; Karon and Saunders come to this conclusion because some unspecified "proponents of heredity" apparently imply that we should act as if there were some necessary relationship between hereditary determination and invulnerability to therapy. No such view has ever been held by either of us and, therefore their whole argument is quite irrelevant to the critique of our experiment. The efficacy or otherwise of psychotherapy and physiotherapy in neurotic disorders must be demonstrated empirically; it cannot be deduced from any "higher principles". In any case, the hypothesis that predisposition to neurotic breakdown is largely inherited does not carry any implications one way or another with respect to the possibilities of cure or prophylaxis.

Karon and Saunders go on to criticize us for faulty reporting; they do this on the ground that 68 pairs of twins were located but only 50 pairs were tested and their results analysed. They imply that 18 cases were discarded and revert again and again to hypothetical reasons for dropping these 18 cases. There is no basis here for any such criticism, however, as our report makes clear that

* This point will be obvious when we consider the kind of evidence required. In the first place, we would need to study unselected or random groups of subjects; by definition the group of people subjected to psychotherapy is heavily weighted in the abnormal, neurotic direction, and can hardly qualify as *random*. In the second place, we must have some criterion to distinguish between hereditary and environmental influences. The often noted similarity between parents and children with respect to intelligence, neuroticism, etc., is clearly not such a criterion as it is equally explicable in terms of hereditary and environmental influences. Karon and Saunders fail to indicate how the practising psychiatrist can acquire such a criterion or use it in his day-to-day work. In the best of my knowledge, no such simple criterion exists. Karon and Saunders' appeal appears to be to prejudice rather than to fact. An analogy might be the appeal to "simply look around" of the supporters of the geocentric and flat-earth theories, when criticizing Copernicus and Galileo; this appeal also is irrelevant as both sets of theories would equally well explain the phenomena obvious to common sense.

https://doi.org/10.1192/bjp.105.438.76 Published online by Cambridge University Press

68 pairs of twins were *located;* not all of them, however, were tested. We settled on 25 identical and 25 fraternal twins at the outset of the experiment; we anticipated that in a number of cases permission to test would not be granted, or that apparatus might break down in the middle of testing. We, therefore, located a larger number of cases in order to have reserves in hand. As it turned out, however, permission was not refused for any of the twins approached and the apparatus did not break down. Consequently, it did not become necessary to call upon any of the remaining 18 twins. The criticism that we eliminated certain cases after testing for some sinister and mysterious reason as implied by Karon and Saunders cannot, therefore, be maintained.

An interesting feature of our study was the fact that as stated in our paper "in 16 out of 18 cases the identical twin variance is larger; two of these differences are significant at the 2 per cent. level". Karon and Saunders use this fact to criticize our calculation of Holzinger's h^2 (11, 10) and present an alternative value which would reduce the hereditary determination of neuroticism to 58 per cent.* The question they raise is an important one. The concept of statistical significance is, of course, an arbitrary one; to say that a difference is significant at the 5 per cent. level simply means that it would have arisen by chance in only one case out of twenty. Many people prefer a higher significance level such as $p = \cdot 01$, i.e. the probability of a given difference having arisen by chance is only one in a hundred. The particular value chosen depends in part on the *a priori* probability of a given event; thus psychologists have quite rightly demanded much higher probabilities in the case of extra-sensory perception where the *a priori* probabilities are very much against the occurrence of such events, than is usual in more orthodox research. Now as Karon and Saunders point out, it would have been easier to understand and rationalize the occurrence of a higher variance for the *fraternal* than for the *identical group*. The opposite finding has a very low *a priori* probability and would, therefore, in my submission require a higher probability value than five per cent. for its substantiation; indeed anything below the one per cent. level would hardly be regarded as more than suggestive. Under the circumstances we felt than on replication of the experiment it would be unlikely that the observed difference would be found again and we decided, therefore, to regard it as accidental. We were confirmed in this decision by the fact that two of the best neuroticism tests (autokinetic movement and suggestibility) had very high h^2 values (·648 and ·701) although variances were not significantly different.

It may be helpful to invert the argument. If we had claimed as a demonstrable fact that variances for identical twins were greater than the variances for fraternal twins (${}_i\sigma^2 > {}_f\sigma^2$), simply because in one study this highly unlikely inequality had been found, we would rightly have been criticised for over-interpretation. Taking all the known facts into account, i.e. refusing to use blindly an arbitrary criterion of "significance", we would conclude that the data did not disprove the null hypothesis (${}_i\sigma^2 = {}_f\sigma^2$) at a level acceptable to us. If on repetition a similar result to ours should be found, we would of course have to revise our conclusion; it is our belief that repetition of the experiment would not duplicate this particular finding.

* They conclude rather mysteriously that "under conditions where much if not most environmental variance is held constant, only about 30 per cent. of a crude neuroticism criterion may be attributed to hereditary determinants". They do not explain how they arrive at this figure, which seems to owe little to mathematical calculation and much to rather unwarranted guesswork. Such speculations do not appear to further the scientific study of genetics as applied to human behaviour.

https://doi.org/10.1192/bjp.105.438.76 Published online by Cambridge University Press

It must be freely admitted that our case rests on an argument which may not appeal to all statisticians; indeed, there is much controversy on this precise point. On reflection it is still my opinion that our conclusion is preferable to that arrived at by Karon and Saunders. It was our hope when originally writing this paper, that others would repeat the experiment with larger samples and better tests than we had available at the time. Until such repetition is in fact carried out, it is impossible to arrive at a decision between the two alternative methods of analysis which cannot be criticized statistically. (It should be noted, of course, that even if we accept Karon and Saunders' figure, we would still find a substantial hereditary contribution to neuroticism; thus even on their showing our main contention would still receive support.)

Karon and Saunders go on to question the validity of our neuroticism criterion. I find it impossible to follow their argument. They quote in a table the biserial correlations between tests and criterion as well as the h^2 of each test, and comment that "we cannot judge from these data whether a better 'neuroticism' factor could be extracted from the battery". This is hardly surprising because the figures in this table by themselves cannot, of course, be used in that fashion. As explained in the article, the factors originally obtained were rotated until the new factor 1 achieved maximum correlation with the criterion column. It is this method of "criterion analysis" (2) which ensures that no better neuroticism factor could be extracted from the battery, and the table quoted by Karon and Saunders is quite irrelevant.

Karon and Saunders make several further comments, some of which are justified, some not. They point out that the differences between the normal twins and the neurotic criterion children "include any differences inherent in *twins versus non-twins* as well as between normals and neurotics". This is true, but not in our opinion important .The point might have had more weight if we had selected our tests simply on the basis of the differentiating power between the twins and the neurotics. By interposing a factor analysis and using the factor score, we have made it much less likely that the twin-non-twin difference would affect our results. The other point made by Karon and Saunders is that the differences between normals and neurotics "are further blurred by the fact that the 'normal' sample of twins must contain some neurotics; indeed if it does not this sample cannot contain variance associable with neuroticism and the whole experimental design collapses". This argument is quite erroneous. It is equivalent to saying that in comparing a group of tall and a group of small people, that the latter group must have contained some giants as otherwise there would be no variance relating to height. Neuroticism is regarded as a quantitative variable and there is no need for the neurotic group to overlap with the normal, although it is, of course, quite likely that in actual fact the groups did overlap. The use of this argument which implies that "neurotics" are a categorical group discontinuous with normals makes me feel that Karon and Saunders have not properly understood the underlying logic of our experiment.

A last point made by Karon and Saunders relates to the selection of tests. Looking at the h^2 values for the tests they say that "these figures provide a striking impression that the tests might just as well have been chosen on the basis of having a high hereditary determination . . . it is quite clear that *any* factor extracted is likely, because of the nature of the tests, to be heavily weighted towards the side of heredity" and in their summary they say that "the validity of the neuroticism criterion itself is . . . biased in favour of

https://doi.org/10.1192/bjp.105.438.76 Published online by Cambridge University Press

hereditary components through the selection of tests". It is difficult to understand the basis of this criticism. The tests were chosen because on the basis of previous work we considered them to be possible measures of neuroticism;* we did not know for any of these tests whether performance on them would or would not be determined to any great degree by heredity. Some of the tests succeeded in measuring neuroticism; others failed. The neuroticism factor extracted from the whole battery segregated neurotic and normal children reasonably well (7), so that we may conclude that our original choice, though far from perfect, was not unreasonable. The fact that the tests did in fact have relatively high h^2 values and consequently showed evidence of strong hereditary determination, may go counter to Karon and Saunders environmentalist beliefs but it is difficult to see how this affects the issue. Indeed, the longer I study the argument the less I understand precisely what they are trying to prove. What they say would only be relevant if we had had available a large battery of tests, the hereditary determination of each of which was already known, and had then with malice aforethought picked out all those tests having high h^2 values while leaving out those with low h^2 values. As we had no *a priori* knowledge of this kind we could not have carried out such a programme even should we have wished to do so.

Having disagreed with almost every point brought forward in criticism by Karon and Saunders, I would like to end by saying how much I agree with their final conclusion. They say "only our efforts to understand determining mechanisms seem likely to produce successful attempts to remedy or prevent situations [*sic*] of any sort including the neuroses. Such understanding of the underlying order in the universe is the proper goal of scientific investigation". I take this to mean that we should try to study in greater detail the precise method of hereditary determination of neuroticism, rather than rest content with a simple numerical estimate of its contribution. Such investigations, making use of the most recent methods of polygenic analysis, are at the moment being carried out in our laboratory, including the method of "diallel crossing". These may succeed in giving a more direct proof of the action of genetic forces, as well as estimates of linkage and dominance. Until these studies are completed it would seem useful, however, to repeat the Eysenck-Prell study with suitable technical improvements, in order to throw some further light on the relative importance of the factors in question. My belief that such studies would support the result of the original paper is increased considerably by the fact that in recent work from the Genetics Unit of the Institute on identical twins brought up in separation, sizable correlations were found for both neuroticism and extraversion, as well as for intelligence (12); indeed, the order of size of the correlations, bearing in mind the reliabilities of the tests, were such that our own finding of similar degrees of hereditary determination for personality traits as for intelligence might be considered supported (4).

*Karon and Saunders, in discussing our selection of tests, complain that the basis for deciding on the particular battery chosen "is nowhere made clear by Eysenck and Prell". We relied on the results of work with children currently proceeding then by Himmelweit and Petrie (8), Connor (1), Thorpe (13) and others, as well as our own experimental studies of neuroticism in adults (2, 3). Karon and Saunders do not suggest how our selection could have been improved on the basis of knowledge available then; at the present moment it is not impossible that a better set of tests could be selected, partly due to recent work on perceptual indicators of neuroticism (5), and partly to the linking-up of learning phenomena with the same personality variable (4). But this knowledge was not then available.

https://doi.org/10.1192/bjp.105.438.76 Published online by Cambridge University Press

References

1. Connor, D. V., "The Effect of Temperamental Traits on Group Intelligence Performance of Children". 1953. London Ph.D. Thesis.
2. Eysenck, H. J., "Criterion Analysis: An Application of the Hypothetico-Deductive Method to Factor Analysis", *Psychol. Rev.*, 1950, **57**, 38-65.
3. Eysenck, H. J., *The Scientific Study of Personality*, 1952. London: Routledge and Kegan Paul.
4. Eysenck, H. J., *Dynamics of Anxiety and Hysteria*, 1957. London: Routledge and Kegan Paul.
5. Eysenck, H. J., Granger, G., and Brengelmann, H., *Perceptual Processes and Mental Disorder*, 1957. London: Chapman and Hall.
6. Eysenck, H. J., and Prell, D. B., "The Inheritance of Neuroticism. An Experimental Study", *J. ment. Sci.*, 1951, **97**, 441-465.
7. Eysenck, H. J., and Prell, D. B., "A Note on the Differentiation of Normal and Neurotic Children by Means of Objective Tests", *J. clin. Psychol.*, 1952, **8**, 202-204.
8. Himmelweit, H., and Petrie, A., "The Measurement of Personality in Children", *Brit. J. educ. Psychol.*, 1951, **21**, 9-29
9. Karon, D. P., and Saunders, D. Q., "Some Implications of the Eysenck-Prell Study of 'The Inheritance of Neuroticism': A Critique", *J. ment. Sci.*, 1958, **104**, 350-358.
10. May, J., "Note on the Assumption Underlying Holzinger's h^2 Statistic", *J. ment. Sci.*, 1951, **97**, 466-467.
11. Newman, H. H., Freeman, F. N., and Holzinger, K. J., *Twins: A Study of Heredity and Environment*, 1937. Chicago. Univ. Press.
12. Shields, J., "Twins Brought Up Apart", *Eugenics Rev.*, 1958, **50**, 115-123.
13. Thorpe, J. G., and James, D. P., "Neuroticism in Children. I. An Investigation of Normal and Neurotic Group Differences", *Brit. J. Psychol.*, 1957, **48**, 26-34.

https://doi.org/10.1192/bjp.105.438.76 Published online by Cambridge University Press