#98

Paul E. Meehl. *Psychodiagnosis: Selected Papers* (pp.200-224). Minneapolis: University of Minnesota Press, 1973. (Paperback, Norton)

12

MAXCOV-HITMAX: A Taxonomic Search Method for Loose Genetic Syndromes

AT THIS WRITING it is no longer possible for an informed person, unless he is an environmentalist fanatic, to believe that everyone is born with an equal biological talent for developing schizophrenia, the only important difference between schizophrenics and others arising from their social learning experiences. Two monumental investigations alone would suffice to make such a view completely untenable (Heston, 1966, 1970; Gottesman and Shields, 1968, 1972). I think it is time for those of us interested in behavioral genetics to suspend debate with radical environmentalists, calmly recognizing that there are ideologies in science (as in politics and religion) which are, for all practical purposes, temporarily resistive to the influence of counterevidence. (See Barber, 1961; Kuhn, 1970; Lakatos and Musgrave, 1970; and the writings of Feyerabend cited on pp. 229-230 therein; but see also Nash, 1963. A fascinating account of the interplay among geneticists between fact and speculation, "rigor" and "looseness," and the fine line between dogmatism and fruitful theoretical tenacity is to be found in Carlson, 1966.) The research task is no longer to find out whether genes have

AUTHOR'S NOTE: This paper is an expansion of an invited lecture delivered at the first convention of the Classification Society (North American Branch), Battelle Memorial Institute, Columbus, Ohio, April 9, 1970. I am indebted to the Carnegie Corporation for support of a summer appointment as professor in the Minnesota Center for Philosophy of Science, and to my colleague Dr. Irving I. Gottesman for suggestions on improving the manuscript.

something to do with schizophrenia, or to convince those who have been brainwashed by American "dynamic psychiatry" and social science doctrine. We should pass on to what is now the scientifically important question, to wit, "Just *what* is inherited and *how* is it transmitted?" (Heston, 1970; Meehl, 1972b—reprinted here as Chapter 11.) It is worth mentioning, however, that determining the mode of inheritance might have an impact even upon the most rabid environmentalist, since the making of successful statistical point-predictions for a syndrome or pathognomonic sign based on, say, a dominant gene theory would constitute pretty strong corroboration. Thus, if the DZ twins, ordinary sibs, and parents of carefully diagnosed schizophrenics showed an incidence $\simeq 1/2$, grandparents, uncles, aunts, and half-sibs $\simeq 1/4$, and cousins $\simeq 1/8$, of a quasi-pathognomonic sign of "schizoidia" (= schizoid tendency, schizoid disposition, schizotypy; see Rado, 1956, 1960; Rado and Daniels, 1956; Meehl, 1962—reprinted here as Chapter 7; Meehl, 1964; Heston, 1970), not only would such a finding argue against a polygenic model (with which it is statistically inconsistent—although the inconsistency may be very hard to detect with unreliable measures, small or moderate size samples, and "unlucky" threshold values, as shown by Edwards, 1960; see also Edwards, 1963, 1969; Falconer, 1965; Murphy, 1964; Dalén, 1969); it would indirectly go against an environmental model, not because such findings are incompatible—they cannot be since the environmental theory generates no point-predictions!— but because in the usual scientific sense a "strong" theory which makes point-predictions is preferable to a weak theory which, while not *refuted* by certain empirical point-values, is incapable of generating them. Putting it another way, since the pure social learning view of schizophrenia does not imply any such point-predictions, a rational man could hardly say that a successful prediction of the point-values 1, 1/2, 1/4,and 1/8 for MZ twins, first-, second-, and third-degree relatives respectively was a sheer coincidence, or that a complex social learning model happens mysteriously to generate precisely the same fractions as flow from a dominant gene hypothesis. (See, for the methodological point involved, Platt, 1964; Lykken, 1968; Meehl, 1967a, 1970a and references, especially to Sir Karl Popper, cited therein.) Problem: Is it somehow possible to generate numerical point-predictions—as contrasted with mere directional significance tests—as to yield "difficult hurdles" and "strong inference" in the Popper-Platt

sense, despite the open-concept (Pap, 1953, 1958b, Chapter 11) status of a loose syndrome?

Clean corroborations would ideally utilize a pathognomonic sign for identifying the schizotype or, lacking that, a set of symptoms each sufficiently strong that a clear identification of the syndrome is possible when they are taken as a group. It seems rather improbable on any plausible theory of schizophrenia that such a pathognomonic sign or sign pattern will be discovered, unless it is biochemical or, possibly, neurological. It can hardly be anticipated that any such sign or sign pattern will be found at the level of molar behavior studied by either psychiatric or psychometric techniques. And if we deal, as we probably will, with so-called "soft" neurological signs—such as experienced clinicians (going back to Kraepelin, Bleuler, and Schilder) have often noted even in the nonpsychotic schizotype—we will be dealing with a loose cluster of highly fallible indicators rather than anything pathognomonic or quasi-pathognomonic.

As I view the current research situation in the genetics of schizophrenia, this constitutes our main methodological hang-up. Most investigators now realize that research relying on formal diagnoses of schizophrenia will probably not enable us to pass much beyond the statement "Genes have a lot to do with this disorder." We would like to substitute a high-confidence diagnostic criterion that an individual relative of a schizophrenic proband is or is not a schizotype. But even that may not be attainable for a while yet. The next best thing would be probability numbers associated with fallible sign patterns, which also—like pathognomonic signs or quasi-infallible syndromes—generate specific point-predictions for a Mendelian model. At the risk of exaggeration, but with the hope of saving some taxpayer money, I would say that very little further research on schizophrenia genetics is likely to be illuminating until a better means of identifying the clinically compensated or semi-compensated schizotype is available. My behavior geneticist colleague Professor Irving Gottesman keeps needling me about my theory of schizophrenia, which was published a decade ago (Meehl, 1962—reprinted here as Chapter 7), saying, "Meehl, you have an interesting theory but your time has run out for testing it." Other than the feeble defense that there ought to be a place for a theoretical psychologist as there is for a theoretical physicist, a more honest reason for my failure to publish empirical evidence for or against my theory is

that *I have not known how to test it*. And I am not interested in adding one more article to the vast and dismal literature of schizophrenia research, most of which, in my opinion, does not tend appreciably to confirm anything (except that this is a complicated disorder, and that psychologists are not very clever about devising "strong inference" methods).

Consider the theory that schizoidia is determined by a dominant gene. I avoid mentioning the penetrance, for reasons which are better given after I explain my statistical method. How might we go about estimating the probability that a particular individual carries the gene, relying on a loose cluster of highly fallible phenotypic indicators? (I use the neutral term "indicator" because not all indicators in the psychological domain are "sick" or pathological enough to be called "symptom," and they are not sufficiently valid to be called "sign" (*Dorland's Medical Dictionary*, 1965; Cronbach and Meehl, 1955-reprinted here as Chapter 1; Meehl and Rosen, 1955—reprinted here as Chapter 2; Meehl, 1959b—reprinted here as Chapter 5). It is imperative, in thinking about the methodological problem, to recognize that there is not presently available any diagnostically definitive touchstone, sign, symptom, or trait which we know how to measure reliably. Even what I view—following Bleuler—as the sine qua non of the disease entity, to wit, thought disorder or "cognitive slippage," will not do for genetic research purposes. There are certain clinical manifestations of cognitive slippage which can be used as quasi-infallible *inclusion* tests, i.e., pathognomonic when present. But these are too deviant to be employed safely as *exclusion* tests, and we do not have any psychometric or clinical device for assessing subtle, subclinical, episodic cognitive slippage of the kind we are accustomed to detect during intensive psychotherapy of schizoid patients, including those that rarely or never show diagnosably psychotic degrees of decompensation. We have therefore a beautiful example of a "bootstraps" problem (Cronbach and Meehl, 1955) in which we start with a fallible set of indicators of unknown relative weights (Meehl, 1959b) and somehow end up assigning weights on the basis of the internal statistical relationships of the elements in this cluster. We have no "acceptable criterion" in the traditional sense. In other words, we have a problem akin to the classical problems of factor analysis, cluster analysis, latent structure analysis, and the like. A schizophrenia theorist or investigator who hasn't reached at least this stage of sophistication

is not in the ball park (I would say he's hardly in the league). It goes without saying (among geneticists) that one of the strongest evidences of a "successful bootstrapsing" operation is resultant conformity of family statistics to a strong genetic model—although one finds psychologists who (with undergraduate canniness) view such arguments as "circular." That the business of a scientific theory is to "carve nature at its joints" is, I trust, not something one must take time defending to professional readers.

How do we decide which indicators belong in the provisional indicator set? We have to decide this on the basis of some combination of clinical experience, previous research evidence—relying upon formal diagnosis merely as a means of "getting our foot in the door"—and, we hope, at least the sketch of a theory. Thus, for example, in attempting to assess the frequency of schizotypy among the relatives of schizophrenic probands, I would certainly include one or more neurological indicators, such as subclinical Romberg sign, a tendency to past pointing, kinesthetic aberration, or paradoxical influence of alcohol upon post-rotatory nystagmus. Why would I see such "soft" neurological signs as good candidates for an indicator set? First, because clinicians have often found such soft neurological signs and transitory subjective neurological "complaints" among patients diagnosed as schizophrenic or schizotypic on other grounds; second, because there is some research evidence to indicate that such signs have validity when the crude criterion is taken as diagnosed schizophrenia; and finally, because my speculative neurological hypothesis quasi-implies that if you are lucky or clever where you look, you *should* find soft neurological signs in schizotypes even when they are not psychotic.

In bootstrapsing fallible indicators of schizotypy there are some special statistical problems which are less likely to arise in non-behavioral genetics. We cannot, for instance, begin our bootstrapsing operation by assigning initial weights on the basis of concurrent validity for diagnosable (psychotic or semipsychotic) schizotypes, because part of the reason we have such a serious diagnostic problem here is that these weights will be very different for the compensated case. It cannot even be excluded that in some instances an indicator might function backwards. For instance, in the "natural history of the disorder" we find clinically that a patient who succeeds in reducing his anxiety by consolidating a paranoid projection system may become more aggressive and extra-

punitive than the average or normal person; whereas, before this defensive resolution, he would have been rated by peers, relatives, and even some professionals as being underaggressive. Nor can we safely assume such familiar statistical approximations as normal distribution or, homogeneity of variances. One of the best established generalizations about the population of clinical schizophrenics is that they are more variable—both longitudinally (over time) and cross-sectionally (in the sense of individual differences)—than controls, a tendency presumably heightened when the population under study includes all schizotypes rather than only the subset decompensated to the point of receiving a formal psychiatric diagnosis of schizophrenia.

Consider a provisional indicator set of not less than three phenotypic variables, deliberately chosen on the basis of the criteria above, plus plausible grounds for hoping that they will be pair-wise uncorrelated (or approximately so) within the postulated latent taxa. So I am making the same assumption as in Lazarsfeld's latent structure analysis, that the observed correlation between the indicators is almost wholly attributable to the influence of the latent taxa. This assumption need not remain an "assumption" in the technical sense of the statistician, i.e., something we postulate without having any means of testing it on the data (as, for instance, the psychometric assumption underlying an arbitrary normalized transformation of test scores). The assumption it-self can be a statistical hypothesis subject to refutation, and I have developed a group of "consistency" tests which should help us decide whether the intrataxon independence assumption is being grossly violated (Meehl, 1965a, section 9; Meehl, 1968, section *g*). Further, we can raise the odds that this independence condition will be approximately fulfilled, at least close enough for the use of the proposed bootstraps method. First, we may rely on theoretical considerations, such as that indicators sampling different behavioral domains or different neurological systems—having, so to speak, very little "qualitative phenotypic similarity or overlap"—ought to be relatively independent. Second, we can ascertain the empirical correlation between the indicators within a group of normals (where the base rate of schizotypy can be safely taken as so low as not to be capable of generating a correlation) as well as among diagnosed schizophrenics, and then extrapolate to the working hypothesis that if a pair of indicators is uncorrelated *within* the schizophrenic group and *within* the normal group, it will

probably not be *markedly* correlated among nonpsychotic schizotypes found in the "normal" population or in a mixed psychiatric population with an erroneous (nonschizoid) diagnosis. Third, in the case of psychometric indicators such as scores on a personality inventory or a "mental status" checklist (Meehl, 1964) or rating scale, we can employ item-analytic procedures to reduce the intrataxon correlation, which we will be willing to do at the expense of sacrificing some amount of validity. I cannot yet make any general statement about the robustness of my method with respect to this assumption of zero intrataxon correlation, although I have some numerical examples (e.g., Meehl, 1965c, section 13) as well as some empirical data on one taxonomic problem, suggesting that a Pearson *r* running up to .30 or .40 may not be too damaging.

An illustrative example of such a provisional indicator set would be (a) a psychometric measure of subclinical cognitive slippage based upon intransitive ("irrational") choice behavior (Braatz, 1970); (b) a measure of the paradoxical effect of alcohol ingestion upon post-rotatory nystagmus (Angyal and Blackman, 1941); (c) a patient's score on a structured personality inventory measuring (by self-report) the phenomenology of pleasure deficit (Rado's *anhedonia*; see Rado, 1956, 1960; Rado and Daniels, 1956). These three kinds of behavior are sufficiently different in the kind and level of dysfunction tapped that one would be surprised to find them appreciably correlated either in a normal population or in a clinical population of nonschizoid psychiatric patients from which organic brain disease, mental deficiency, and grossly psychotic cases had been excluded. So we have here three tentative indicators of schizotypy which we have plausible reasons to hope are relatively independent except as they are influenced by the hypothesized dominant schizogene.

I want to emphasize how little we know by way of commencing our bootstraps operation even if we assume the above. We do not know the relative validity of these three indicator variables, and we cannot estimate it by relying on cases of diagnosed schizophrenia. Putting this more generally, we do not know what the means and variances, or even the distribution forms, of the indicator variables are within the postulated latent taxa. We cannot safely assume that the distributions are homogeneous in variance, or that they are normal. (As a matter of fact, if we extrapolate from research on diagnosed schizophrenics there

is good reason to think they will not satisfy either of these assumptions.) Finally, we do not know the base rate *P* of schizotypy in a population at high risk, such as the first-degree relatives of diagnosed schizophrenic probands, or in a mixed psychiatric population. For readers unfamiliar with the diagnostic situation in psychiatry I should mention that not only is this very important parameter *P* unknown, but we cannot even begin with a plausible estimate of it for bootstrapsing purposes. This is because competent and seasoned clinicians, holding differing views about schizophrenia, assign markedly different indicator weights to the clinical phenomena. One can find boarded psychiatrists or clinical psychologists who, when asked, "What is the base rate *P* of schizophrenia (or latent schizophrenia, or subclinical schizophrenia) in a general outpatient psychiatric population?" will give estimates ranging from a low of 10 percent—particularly if they are British or Continental psychiatrists whose conception of schizophrenia is rather close to Kraepelin's dementia praecox—to a high of 90 percent (I have actually heard this figure from a very capable psychiatrist trained by Rado). And nobody is presently in a position to refute either of these extreme values. So we start out with very little tentative knowledge. and a large amount of ignorance. It might seem impossible to get anywhere bootstrapsing from such feeble foundations, but unless I have made a mistake, I think we can do it. The very inadequacy of our antecedent information is reassuring, because it frees us of the obligation to show that our proposed bootstrapsing method is highly precise, that it leads to maximum likelihood estimates, that the sampling errors are very small, or to provide analytical derivations of random sampling distribution functions for the consistency tests. The point is that when you know this little, even a rather crude method, so long as it seems to check out on real data (as in the example I shall present later) and to be reasonably robust on Monte Carlo study, may legitimately be employed. I may say that I find some statisticians puzzling in this respect, because they seem to me to be saying, in effect, that if a proposed method does not lead to elegant mathematical answers or to studentized sampling distributions derivable analytically, the method isn't even worth exploring. But meanwhile they themselves seem often to be making assumptions about the latent situation or the state of prior knowledge which are unrealistic in the behavior genetics of psychoses and neuroses. My view is that one is better off with approximate methods that are realistic in the research

context *and help to answer an important question* than with "precise" methods for which the assumptions and input information are not fulfilled, or which answer uninteresting questions, e.g., "Is the null hypothesis false?" (See, in this connection, Badia, Haber, and Runyon, 1970; Morrison and Henkel, 1970.) When I first presented this paper at a meeting of the newly formed Classification Society, I experienced trepidation over the fact that I had no analytical derivations of exact sampling distributions for the statistics proposed. It was surprising and reassuring to discover that the mathematicians, statisticians, and computer experts who there presented new taxonomic search methods rarely had any such either, and were strangely freewheeling about it.

So we have three quantitative variables x, y, and z of unknown relative weight, of unknown variance and distribution function, and we do not even have a base rate (in this context, gene frequency) for the taxon of interest. We hypothesize that each indicator has some moderate to high construct validity (Cronbach and Meehl, 1955—see Chapter 1 above; Campbell and Fiske, 1959; Loevinger, 1957) and that the indicators are uncorrelated pair-wise within the schizoid taxon and outside it. As to distribution form, we hypothesize that while perhaps skew or leptokurtic or platykurtic, each indicator variable is at least unimodal within the two groups. I note in passing that even this weak assumption may actually be false for schizotypy, there being some evidence to suggest a bimodality when we mix schizophrenics of the paranoid and nonparanoid subgroups.

Consider one of the three indicators z, which I shall call, for reasons which will be apparent in a moment, the "input indicator." This does not mean "input" in the causal sense, but simply input in the context of our search technique. We therefore imagine the latent situation, the state of nature as known to Omniscient Jones but not to the investigator, as in Figure 1. These are unrelativized frequency functions rather than probability-density functions, so the ordinates and areas reflect the different base rates P,Q.

Suppose the clinician or researcher, being ignorant as he is of the parameters of these latent frequency functions, draws an arbitrary cut on the abscissa, dividing the manifest (mixed taxa) empirical distribution, and labels patients falling above this cut as "indicator positive." The area under the upper distribution lying above this cut represents the "valid positives," that is, the cases classified by the cut as schizo-

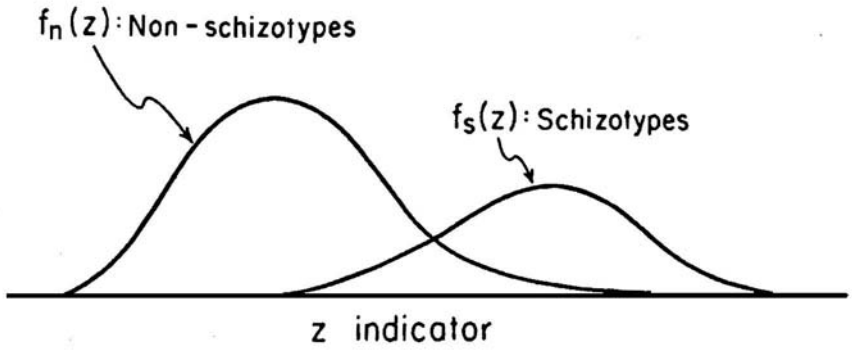

Figure 1. The latent situation

types who are in fact members of the schizotypal latent taxon. The *proportion* of the schizotypal distribution falling above that cut we label the "valid positive rate," symbolized by p_s. Note that this rate is not the proportion of indicator positives that are in fact schizotypes, but the proportion of true schizotypes who are correctly identified as such by the cut, i.e., the denominator of this rate p_s is the true schizotype base frequency N_s. Similarly the proportion of cases in the lower (non-schizotype) distribution lying above that cut we designate as the "false positive rate," symbolized by p_n.

For any such fallible indicator, i.e., any indicator in which the two indicator functions overlap, shifting the cut results in an improvement in one of these rates at the expense (worsening) of the other. By moving the cut downward we increase the proportion of schizotypes correctly so labeled (i.e., p_s rises) for which we pay the price of an increase in the false positive rate p_n. If we wish to reduce the false positive rate p_n, we have to move the cut upward, which identifies fewer of the true schizotypes and therefore gives us a reduction in the valid positive rate p_s. The "optimal" cut depends therefore not only upon the character of the two probability functions but upon the base rate P (proportion of true schizotypes in the mixed population under study). For a discussion of the practical clinical problems arising from this state of affairs see Meehl and Rosen (1955—Chapter 2 above). For purely research purposes, as in testing a dominant-gene hypothesis, the optimal cut is one that minimizes the misclassifications; for clinical purposes, however, it may not be total misclassifications we desire to minimize but mis-

classifications weighted by the clinician's assignment of disutilities attached to the two kinds of errors.

It is worth noting that from the psychologist's viewpoint, accustomed as he is to moderate or low validities involving a sizable overlap of indicator functions, the geneticist's concept of "penetrance" (taken here in the sense of a single phenotypic indicator variable) suffers from arbitrariness. The penetrance coefficient of a dominant gene in this situation corresponds to the valid positive rate p_s, and to the concept called *sensitivity* in epidemiology. There is, so far as I know, no standard term in genetics for designating the false positive rate p_n, the complement of which, $(1 - p_n)$, is labeled *specificity* by the epidemiologist. From the psychometric standpoint, penetrance is a derivative concept, the geneticist's *expressivity* being the fundamental one. That is to say, what I would call the *expressivity function* is nonarbitrary, being a mathematical fact about the state of nature for a given genome-cum-environment joint distribution characterizing a specified population. We may not know at a given stage of research what that function is, but we know that such a function exists in the state of nature and is not arbitrary, whereas penetrance, except in the case where the distributions are nonoverlapping (penetrance of 100 percent), is an arbitrary function of the sliding cut. By increasing the proportion of false positives, i.e., by lowering the cut, we increase the penetrance of the gene with respect to a given indicator variable. It is my understanding that in the case of fallible indicators, what geneticists tend to do is to choose a cutting score more or less arbitrarily rather than to optimize the cut (like an industrial psychologist) in the light of the base rate P. For example, in considering the palm lines associated with Mongolism, a cutting score of 57° in maximum atd angle is set as aberrant, following Penrose (1954). In our notation, this cut on a quantitative indicator variable yields $p_s = .80$ for Mongols, $p_n = .07$ to .09 for the general population, and $p_n = .14$ to .16 for mothers and sibs of Mongols (Stern, 1960, pp. 471–472). (Are these last all "false positives," and, if so, why do they run almost double the general population (+) rate?) Strictly speaking, the optimality of a cut at 57° would depend upon these parameters, the base rate P of the Mongol karyotype, and the research or clinical context. One may of course choose some suitably low value of false positives and treat the errors as essentially negligible, such as the 5 percent or 1 percent point on the lower frequency function; this is satisfactory

for most clinical and research purposes in other branches of pathological genetics, where we typically deal with the combination of high validity indicators and minuscule base rate *P*. But in research on such a loose syndrome as schizotypy, and one with a high gene frequency, a less arbitrary procedure is desirable.

Calling a "hit" a schizotype falling above the cut or a nonschizotype failing below the cut, the total correct classifications (taking account of the base rate *P*, i.e., the gene frequency in a high-risk population) is given by

[1] $H_t = H_n + H_s = F_n(z) + [N_s - F_s(z)]$, where $F_n(z) = \int_{-\infty}^{z} f_n(z)dz; F_s(z) = \int_{-\infty}^{z} f_s(z)dz$

and to maximize the total hits yielded by cutting a single-indicator variable we set the derivative of this sum at zero

[2] $F'_n(z) - F'_s(z) = 0$

[3] $f_s(z) = f_n(z)$

which means that the optimal cut for minimizing total errors occurs at the abscissa value below the intersection of the two frequency functions, i.e., where the ordinates f_s and f_n are equal. This is intuitively obvious from the geometry of Figure 1. I shall designate that abscissa value the *hitmax cut on z*. But of course we do not know the latent frequency functions f_s ,f_n. Can we locate the hitmax cut on z by studying the behavior of the other two indicators (x, y)? Intuitively, on our provisional assumption that the variables of the indicator set are pairwise uncorrelated within taxa, it is obvious that any observed correlation—I am neglecting sampling error throughout this paper—will be attributable to the existence of taxon mixture. That is, if we had a subpopulation consisting wholly of schizotypes, or one consisting wholly of nonschizotypes, the correlation (or, as will be more convenient to work with, the covariance) of an indicator pair would be zero. If we were to examine various subpopulations composed of varying mixtures of the two latent taxa, it is intuitively obvious that the observed covariance of two indicators will increase with the amount of taxon mixture, and will be a maximum when the taxon mixture is a maximum, i.e., in the subpopulation composed equally of schizotypes and non-schizotypes. Algebraically, we write the general expression for the covariance of a

mixed population for two indicator variables x and y, where p = proportion of schizotypes, $q = 1 - p$,

[4] $$cov(xy) = pcov_s(xy) + qcov_n(xy) + pq(\bar{x}_s - \bar{x}_n)(\bar{y}_s - \bar{y}_n)$$

The components of this total mixed-taxon manifest covariance *cov(xy)* are then the weighted intrataxon covariances plus a term whose size depends upon the "validities" of the two indicators, represented by the differences of the latent means, and the product pq representing the amount of mixture. On the assumption of zero intrataxon covariance, the first two terms drop out and the expression for the observed covariance of x and y reduces to

[5] $$cov(xy) = pq(\bar{x}_s - \bar{x}_n)(\bar{y}_s - \bar{y}_n)$$

If there were some way to arrange a series of subpopulations beginning with a "pure" population composed solely of non-schizotypes and running through a series of subpopulations in which the proportion of schizotypes increases steadily until we reach a value $p = 1/2$ and thereafter a series of populations in which the proportion of schizotypes increases beyond 1/2 until we get to a subpopulation which is also "pure," consisting only of schizotypes, the manifest (xy) covariance would be seen to begin at zero, increase to a maximum, then to decline again to zero. Taking the product of the latent means $(\bar{x}_s - \bar{x}_n)(\bar{y}_s - \bar{y}_n) = K$ as fixed—ignoring sampling fluctuations—the quantity $cov(xy) = Kpq$ will be maximized for the subpopulation in which the taxa are equally represented, that is, where $p = q = 1/2$. That the product pq is greatest for diagnostic symmetry ($p = q$) corresponds to the intuitive notion that if two indicator variables are correlated solely because of the latent taxa, they will correlate most when the population is "most mixed" taxonomically.

But of course we do have a way of ordering a series of subpopulations in this fashion, namely, we can order them on the basis of our third indicator variable z. Since if one considers the sequence of class intervals on z arranged by taking successive slabs of patients on the mixed frequency function $f_t(z) = [f_s(z) + f_n(z)]$of Figure 1, at the low end of this distribution all of the cases in the class intervals of z are non-schizotypes; at the upper tail all of them are schizotypes; and in the middle we have varying amounts of mixture, the greatest mixture occurring in the interval surrounding the hitmax cut on z. So our search

procedure, taking indicator z as the input variable and the indicator pair (*x,y*) as output variables, consists simply in calculating the (*xy*) covariance for each *z* interval from low to high, and looking for its maximum. Since this is the core idea of the method, I have tentatively christened it "MAXCOV-HITMAX."

With a respectable sample, the orderly behavior of the (*xy*) covariance as a function of the position of each subsample on the *z* indicator tends of course to corroborate the postulated latent model. One can employ some kind of moving average, or—as my research assistant did in an effort to improve the method—fit a function (he fitted a parabola) to the plot of the covariances, although (strangely enough) this determination of the hitmax cut by finding the analytic maximum of a fitted curve did not improve validity.

Locating the hitmax cut on *z* is intrinsically useful, especially since it can be checked by an independent method that relies, however, on a somewhat stronger model, postulating approximate intrataxon normality (Meehl, Lykken, Burdick, and Schoener, 1969). But in locating the hitmax cut on *z* by maximizing the (*xy*) covariance, we have meanwhile obtained the latter's numerical value, and this permits us to make a further inference which is powerful for our bootstraps operation. Taking the product of the latent mean differences as a constant, that is $K = \left(\overline{x}_s - \overline{x}_n\right)\left(\overline{y}_s - \overline{y}_n\right)$, we have in the hitmax interval on *z* the relation

[6] $cov_{hz}(xy) = Kp_hq_h = 1/4\ K$

and since the quantity $cov_{hz}(xy)$ is an observed value (that is, the mixed-taxon (*xy*) covariance obtained on the cases lying in the hitmax interval of *z*), we solve for the product of the latent means on the output indicators = *K*.

Knowing *K*, since on the assumption of zero intrataxon (*xy*) covariance the relation of equation 5 holds within each of the *z* intervals, we can write the general expression for any *z* interval,

[7] $Kp^2{}_z - Kp_s + cov_s(xy) = 0$

a quadratic in the variable *p* = proportion of schizotypes in that *z* interval. For each of the *z* intervals we can plug in the observed (*xy*) covariance for that *z* interval and solve the resulting quadratic for *p*. Multiplying p_i in each *z* interval by the observed frequency n_i for that interval gives us an estimate of the latent frequency of schizotypes n_{si}

within the interval, so that in effect the series of solutions of this quadratic over the z range draws us the unknown latent frequency functions $f_s(z)$ and $f_n(z)$. Summing the values n_{si} and n_{ni} over all z intervals gives us the latent total taxon frequencies N_s and N_n, and dividing these by our total N yields the unknown base rates P and $Q = 1 - P$.

We can then choose a different indicator of the set, say y, and repeat the process above using y as the input variable and $cov(xz)$ as the output variable; similarly we can choose the remaining indicator x as input and plot $cov(yz)$ as output. Agreement between the results of these procedures (on both the base rates obtained and the latent means) provides consistency tests for the adequacy of our idealization.

Suppose we conclude on this basis that the postulated latent taxonomic model is reasonably satisfied and our estimates are consistent enough to be relied upon as a bootstrapsed approximation. We have determined hitmax cuts on each of the three indicators of the set, and for each indicator cut we have estimated the valid and false positive rates P_s and P_n characteristic of the population under study, keeping in mind that these hitmax cuts and the resulting hit rates are not invariant over clinical populations having different base rates. We can now employ Bayes' Theorem to calculate the inverse probability that a patient belongs to the schizotypal taxon. That is, consider a patient who falls above the hitmax cut on indicators x and y but below the hitmax cut on indicator z. What is the probability that he is a schizotype? We write

$$[8] \qquad p(S_c / x^+ y^+ z^-) = \frac{P p_{sx} p_{sy} q_{sz}}{P p_{sx} p_{sy} q_{sz} + Q p_{nx} p_{ny} q_{nz}}$$

so that for each of the eight possible sign patterns (+ + +), (+ + –), (+ – +),…(– – +), (– – –), there is a Bayes' Theorem probability computable for patients showing that specified sign pattern. Even if each of the three indicators taken singly has only moderate validity (corresponding to "low penetrances" of the sort that make geneticists skittish about invoking the concept) we can subject a Mendelian hypothesis such as dominance to a fairly rigorous empirical test because, while we do not know with high confidence for each individual patient, or each relative of a known schizophrenic proband, whether *he* is or is not a schizotype, we can assign probability values to this taxonomic classification; and this possibility leads to the generation of point-predictions for various sign patterns arising in first-, second-, and third-degree relatives of schizophrenic probands.

Not being a mathematical statistician I have not attempted to derive random-sampling distributions of these statistics analytically, and I am informed by my local experts that it would not be possible to do much along these lines without imposing greater constraints upon the latent model than I wish to do. I am therefore engaged in a large-scale Monte Carlo investigation of this question. Unfortunately my only sufficiently large mass of empirical psychometric data (our Minnesota files on MMPI records) is currently unavailable to me, being in the process of careful diagnostic screening and rescoring before being put on computer tapes. But I have conducted one empirical investigation of the method on real data, employing a genetic problem that is known to be taxonomic and where we have an infallible criterion, to wit, biological sex, and working with psychological indicators very remote in the causal chain from the XX and XY genotypes. Taking three highly fallible psychometric indicators of sex, consisting of three item-analyzed masculinity-femininity scales derived from the MMPI item pool, and pretending that we do not know the indicator functions or the base rates, we applied the method on a sample of 1105 psychiatric patients with a true male base rate $P = .39$, which yielded an estimated base rate $P' = .36$, gratifyingly close to the true value; and application of Bayes' Theorem to the eight sign patterns—classifying each patient as male or female depending upon whether the inverse probability of taxon membership was greater or less than 1/2 yielded 85 percent hits.

I have developed about a half-dozen alternative methods of locating the hitmax interval and some nine consistency tests of the latent model. As an example of one of these consistency tests, suppose we apply equation 5 to the *total* manifest distribution, plugging in the estimated base rates and the estimated latent means for each of our three indicator pairs. Then the three grand covariances can be calculated relying on the estimated latent quantities as follows:

[9] $$cov_t(xy) = PQ(\bar{x}_s - \bar{x}_n)(\bar{y}_s - \bar{y}_n)$$

[10] $$cov_t(xz) = PQ(\bar{x}_s - \bar{x}_n)(\bar{z}_s - \bar{z}_n)$$

[11] $$cov_t(yz) = PQ(\bar{y}_s - \bar{y}_n)(\bar{z}_s - \bar{z}_n)$$

It may also be possible to work with situations in which the intrataxon covariances depart from zero too much to take this as an adequate approximation. (So far as locating the hitmax cut is concerned, equation 4 shows that a weaker assumption suffices, namely, that the intrataxon

covariances are at least equal, if not zero. But the next step, estimating *K*, cannot be taken on this basis.) By beginning with the zero intrataxon covariance assumption and making initial estimates of the latent parameters, we draw an arbitrary cut (say, at the median) on each indicator, write the complete equation 4 for the case of non-zero intrataxon covariance, solve these two equations (one based upon the observed subsample lying above the median cut and the other on the cases lying below it) for the two intrataxon covariances as unknowns, plug these values into the grand equation, and recycle until the system settles down. Our results on the sex-classification data do not, perhaps surprisingly, indicate that this iterative procedure improves validity appreciably. I have done some rough paper-and-pencil computations which suggest that the method may be fairly robust under departures from the assumption of zero intrataxon covariance, but this question is obviously in need of thorough Monte Carlo or analytic investigation.

A major limitation of the method is the sizable sample required; but I should point out that the validities of new indicators can be estimated on considerably smaller samples, once we have obtained estimates on old indicators (such as personality test data) available in larger numbers from clinical files. In fact, I have shown elsewhere (Meehl, 1965c; see also Dawes and Meehl, 1966—reprinted here as Chapter 8; Dawes, 1967; for a criticism, see Alf and Abrahams, 1967; an improvement which amounts to a consistency test is given by Linn, 1967) that if we were so clever or lucky as to hit upon a neurological or biochemical indicator *v* that was quasi-infallible in tracing the schizogene, the fact of its quasi-infallibility could be inferred with confidence. We could do this by showing that there exists an optimal cut on *v* such that the observed rates of cut positives $p^{+}(v)$ could be made to conform to the Bayes' Theorem-estimated schizotype rates in the cells of a table defined by the sign patterns on psychometric indicators (*x*, *y*, *z*) of only moderate validity. This result, which I call (because of its paradoxical character) the "Super-Bootstraps Theorem," permits us to begin with file data on indicators of only moderate validity, such as personality test scores or psychiatric behavior ratings. Such indicators are many steps removed (in the causal chain of polygenic and environmental factors) from the gene of interest, but the statistics enable us to locate and validate neurological and biochemical indicators which, being much closer to the gene action, manifest a much higher "penetrance." Numerical example:

Suppose the base rate of schizotypy in a mixed psychiatric population to have been estimated at $P = .40$ employing the MAXCOV-HITMAX method, with hitmax cuts on each of three indicators yielding symmetrical hit rates $p_s = q_n = .70$, which is pretty fallible but presently achievable (e.g., the MMPI schizoid scale 8 has a concurrent validity better than this, against fallible diagnosis as criterion). The eight sign patterns would then provide Bayes' Theorem estimates for latent schizotype rates ranging from $p(S_c/x^+y^+z^+) \cong .83$ to $p(S_c/x^-y^-z^-) \cong .05$ over the eight cells of an inverse, probability table. A new indicator v can now be cut at arbitrary values and the discrepancy of the (v^+) rates from the Bayes' rates per cell tabulated. We choose the cut v_c that minimizes these cell-value disparities. If that optimal v cut achieves a very close fit (ideally, a fit within sampling and psychometric errors) we conclude that v_c is quasi-pathognomonic. Short of such good luck, however, we may be able to infer that the construct validity of v when optimally cut is extremely high—better than that of bootstraps indicators x, y, z singly or jointly.

Not being a geneticist, I am properly hesitant to suggest modifications in genetic terminology, although I have been so rash as to employ the term "potentiator" for designating any of an open class of (presumably polygenic) variables that, in my theory of schizophrenia, increase the probability of a schizotype's decompensating to the extent that he becomes clinically diagnosable as "schizophrenic." My Minnesota colleagues have tried to convince me that the available terminology (e.g., "epistasis," "modifier") suffices to cover what I label "potentiators," and I have no wish to clutter up the language of behavior genetics by a superfluous neologism. But I should perhaps say a few words about why I hesitate to employ the received terminology in expositing my own theory. Perchance such an explanation, even if my terminological proposal is deemed unnecessary, may highlight some methodological issues that have not as yet received sufficient attention. In what follows, I presuppose that any theory of schizophrenia possessing respectable verisimilitude (Popper, 1959, 1962; Lakatos, 1968) will be *at least* as complex, causally and statistically, as that shown in a diagram I prepared for another paper (Meehl, 1972c; see page 190 above). (As a clinician, I find it quite impossible to suppose that temperamental parameters of anxiety, rage, social introversion, dominance, sexual constitution, energy, and the like—all of which have heritable components in humans

as well as infrahuman mammals—should be *irrelevant* to whether a schizotypal individual remains clinically compensated. I am therefore puzzled by those psychologists who find the increased incidence of *non*-schizoid psychopathology among the relatives of schizophrenics a big surprise, suggestive of old-fashioned "neuropathic diathesis" concepts. *Psychodynamically, how could it be otherwise?*) It is this causal complexity that gives rise to the taxonomic search problem discussed in the present paper. Every time we add another link in the causal chain, whether that link is a non-schizo-specific polygenic influence (e.g., anxiety proneness) or an environmental parameter (such as schizophrenogenic mother, lower social class, or a cruel husband) we lower the probability linkage between the postulated dominant schizogene and the behavioral indicator relied on for diagnosis. Furthermore, we are almost certainly dealing with such causally complicated relations as (a) correlated initial and boundary conditions, (b) subject-selected learning experience, (c) social feedback loops, (d) autocatalytic psychological processes, (e) critical junctures in "divergent" causality (Langmuir, 1943; London, 1946), and (f) intrinsically unpredictable "contingency factors" (Horst, 1941). For a methodological discussion of these see Meehl (1970a).

The same semantic doubts that generate my reluctance to employ the geneticist's standard term *modifier* for fear of mishandling it in a situation as complicated as schizophrenia also lead me to wonder whether the word *penetrance* is appropriate. This latter issue is important because one finds that geneticists are troubled by a causal model which would lead to a rather low penetrance coefficient for clinical schizophrenia taken as the phenotypic expression of a dominant gene. My hunch is that we are somewhat misled by taking as our paradigm a neurological disorder such as Huntington's Disease or any of the large number of Mendelizing forms of mental deficiency, where one conceives of the clinical *disease* entity as, so to speak, the "expected outcome" of the genotype of interest, and its absence in some individuals as a kind of "exception" which is a candidate for special explanation. We know, for instance, that in order to fit a strict dominant gene model for Huntington's Disease, one must extend the risk period up to age seventy or more, and we do not get fully 100 percent penetrance even then. I suggest that these paradigm cases lead us astray when we think about schizophrenia, and we should rather think of the "usual," "typical," "to-be-expected"

result of the schizogene's presence as being the subtle, subclinical, non-social neurological syndrome I postulate and label "schizotaxia." We then think of the individual with that fundamental neurological disposition as acquiring, by a complicated process of social learning, the personality structure, dynamics, and mental content that Rado designates "schizotypal." Finally, we postulate that some unknown propor-tion (considerably less than one-half) of schizotypes decompensate to the point that they *would* be clinically diagnosable under careful study; and only a fraction of these latter are actually formally diagnosed, so that they show up in the files available to most investigators (Heston, 1970).

To see how serious this problem is, I may perhaps be permitted to invoke my own clinical experience in the private practice of psycho-therapy. Using an earlier form of the present taxonomic search method, combining the MMPI with therapy-based judgments quantified in my Schizotypal Checklist (Meehl, 1964), I concluded that a little less than half of the therapy patients I had carried over a ten-year period were schizotypal, amounting to some two dozen in number. For many of these I had very high confidence in the schizotypal diagnosis. For quite a few of them there had occurred, in the course of intensive psychotherapy, micropsychotic episodes (Hoch and Polatin, 1949) which, *had they come to the attention of any clinician*, would have led him to agree with me that the patient was (at least transitorily) "schizophrenic," whatever might be our disagreements upon etiological theory. For example, most of these patients had experienced episodes of body-image aberration, hallucinatory and delusional phenomena, severe thought disorder, and the like. The point of this story for present purposes is that at the time I surveyed my clinical files, not one of these two dozen (almost certainly schizoid) patients had ever been formally diagnosed as schizophrenic, or even as schizoid, in any clinic or hospital (cf. Peterson, 1954). Hence, if they had been under statistical study as the relatives of some officially labeled schizophrenic proband, they would have all been discordant! (Two of them have since been hospitalized and diagnosed schizophrenic, however.) This argument may appear to the reader to prove too much. However, these patients were selected by me as sufficiently intact (and in several cases, thereby misdiagnosed) to be suitable for outpatient treatment by myself as a nonphysician. They are therefore a selected sample; but the main point of the story remains, to

wit, that here we have two dozen people who were psychometrically and clinically schizotypal, and the majority of whom had experienced transitory acute schizophrenic phenomena, but who were not down on any medical, educational, or social agency's record as "schizophrenic."

Since I am a neo-Popperian in my philosophy of science, I do not mind going out on a limb in the absence of adequate "inductive evidence"; so I will record my prophecy that the "clinical penetrance" of the dominant schizogene is less than 20 percent, and possibly as low as 10 percent. (I am well aware that some geneticists consider such penetrance values to be methodologically sinful; and that is one reason why I do not like to use the word here.) Taking as a rough base rate of diagnosed schizophrenia the usual figure of around 1 percent, this means that I am taking the prevalence of schizotypes, who carry the dominant gene, to be as high as 5 percent of the population. This is a pretty steep figure if we take Mendelizing "mental defects" as our model; but I remind my geneticist friends that diagnosed schizophrenia itself has a *huge* incidence by their usual standards (as compared to disorders like Huntington's, Turner's, Tay-Sachs', PKU, and so forth). I guess I am saying, as a confessed nonexpert speaking to the experts, "Perhaps we should not be too surprised if this entity turns out to have some quantitative oddities; we already have reason to see it as kind of special." I have elsewhere suggested that the schizoid disposition is more analogous to something like ordinary (and very common) red-green color blindness, i.e., a kind of capacity defect whose "psychological" consequences develop as a result of complex social learning processes (Meehl, 1959b—reprinted here as Chapter 5; 1962—reprinted here as Chapter 7; 1972c—reprinted here as Chapter 11).

I can illustrate my reluctance about whether the terms *modifier* and *penetrance* would be used here sufficiently like their traditional meaning by an analogy which some might think is farfetched, but which I view as misleading only because it is too simple! Consider diabetes, which is admittedly a genetic disorder (although its mode of inheritance is still in dispute). Suppose that a particular carrier of the diabetic genotype has also inherited a heavy loading of anxiety-parameter polygenes, as well as the gene or genes that (on some views) are relatively specific for alcoholism (i.e., that lead the individual to experience that very powerful reinforcing effect of the alcohol molecule, and therefore tend to lead

to overconsumption and addiction, even if the subject is otherwise not under any unusual degree of psychological or social stress). As a result of. his alcoholic tendencies, the patient maintains poor foot hygiene. Also, one night when he is intoxicated, he traumatizes his big toe. He doesn't pay any attention to this, and as a result he develops diabetic gangrene, a common phenomenon but one observed in only a small minority of younger, well-controlled diabetics. The gangrene necessitates that his leg be amputated. This be unconsciously construes, owing to an intense unresolved Oedipus complex (having nothing at all to do with his diabetic tendencies, but related to his anxiety proneness and his hysteroid seductive mother) as a major castration experience. As a re-sult of this symbolic castration, he becomes deeply depressed. Now as a clinician interested in behavior genetics, I do not find anything the least, bit farfetched or implausible about that causal chain. But surely it would be a misleading use of the geneticists' language to say that the genes for anxiety proneness or alcoholic addictiveness—in spite of the very important role they played in the development of this particular patient's troubles—were "modifiers" or "epistatic" with respect to the diabetic gene! And it would also seem rather strange use of the geneticist's terminology to count the percentage of individuals in whom such a sequence had happened, and then to label that numerical value the "penetrance." If the patient had a seductive mother and a terrifying father, which combined to yield such a strong Oedipus complex and, hence, such an exaggerated susceptibility to the castrative meaning of a leg amputation, I cannot imagine that a geneticist would want to refer to those environmental factors as influencing the "expressivity" of the diabetic genotype, would he?

If I had some competence in genetics, I would examine a collection of medical and nonmedical examples involving complicated feedback loops and effects upon behavior, with an eye to formulating some sort of methodological distinction between the kind of polygenic influence that sticks rather closely to the original meaning of *modifier*, as contrasted with the kind of situation I have just described, for which I have employed the word *potentiator*. For instance, in the case of schizophrenia, suppose one holds—as I do—that some of the "soft" neurological signs are fairly close in the causal chain to the schizogene, and that there may be some perceptual and psychometric signs, unfortu-

nately also "soft" statistically, that are almost equally close. Polygenic systems affecting microstructural features of the CNS and altering the probability that a schizotaxic individual would show a particular "soft" neurological sign (e.g., subclinical Romberg) would seem to be appropriately labeled *modifier*. Similarly, if there are genetic determiners for some of the perceptual-cognitive differences that psychologists study (e.g., the augmenter-reducer dimension in the kinesthetic aftereffect) it would not seem an overly stretched usage to apply the term *modifier*. But when we get to such a causally remote link in the chain as whether a patient projects hostile delusional material, and explain this partly on the basis that he is a mesomorph and partly on the basis that he has inherited a high rage parameter (completely independent genetically of the schizogene), one begins to feel that "modifier" is no longer the right word. We have to face the fact that clinical schizophrenia, except in its subtle and still-disputed neurological features, is a collection of socially learned behaviors. We do not need any further research to be able to say that. We need only ask, "What are the behavior dispositions sampled in deciding that a patient is schizophrenic?" They are, without exception, learned social responses, having a learned motivational, affective, and cognitive *content*. And this is just not the sort of thing involved in something like Huntington's Disease. As Bleuler (1911 as reprinted 1950) said sixty years ago in his classic work, a person cannot have a delusion concerning Jesuits unless he has learned about Jesuits. In that sense, the psychodynamically oriented clinician who insists that schizophrenia is not biologically inherited but socially learned is obviously (but not illuminatingly) correct.

Addendum (July 1972)

Subsequent to the presentation of this paper at the Classification Society's meeting, the large-scale Monte Carlo runs then projected have been completed and partially analyzed for major trends of prime interest. Detailed results will be reported elsewhere (see Golden and Meehl, 1973a, 1973b, and papers now in preparation for submission to a psychometric or statistical journal). To avoid further delay in publishing the present volume, I here confine myself to summarizing our main interpretations of the results to date:

1. The MAXCOV-HITMAX method yields highly accurate estimates of

the unknown latent taxon's base rate *P*, typically with an error about that of the standard error of a proportion for each sample size.

2. Joint satisfaction of the three or four best consistency tests practically assures an adequate numerical approximation by the state of nature to the idealized latent structure.

3. The method is gratifyingly robust, leading to accurate parameter estimates even when intrataxon indicator correlations depart markedly from the idealized $r = 0$.

4. Only a "malicious" combination of highly adverse latent circumstances (e.g., extreme inequality of intrataxon variances, marked indicator skewness, base rate asymmetry $P \ll Q$, small sample size, large and unequal intrataxon covariances) seems capable of yielding bad parameter estimates *without clear warning* (*by consistency tests*) *that this is happening*. Even for these far-out "bad luck" combinations, we have reason to conjecture that an optimal configural use of the consistency tests may yet succeed in reducing the inferential danger to a negligible threat.

5. Application of the method to some further "real data" (U.K. psychiatric ratings) strongly suggests that it works much better than the leading "cluster method" contender.

6. A related but nonredundant method, MAXDIFF-HITMAX, which locates the hitmax cut on indicator *x* by searching for the maximum difference between means of indicator *y* calculated above and below a sliding *x* cut (i.e., we assign x_c such that $[\overline{y}_{(x>x_c)} - \overline{y}_{(x<x_c)}] = \text{Max}$), is as good as MAXCOV-HITMAX, perhaps better. This procedure (see Meehl, 1965a, section 9d, pp. 29–30; Meehl, 1968, section 2b, pp. 9–23) can be powerfully combined with an "item-iterative" approach, in which the taxonomic parameters p_s and p_n of individual MMPI items are estimated in a first approximation, then used to improve the hitmax cut location and to eliminate "bad-acting" items, and continuing to recycle until consistency tests are well satisfied by the surviving items. Since MAXCOV-HITMAX and MAXDIFF-HITMAX rely on largely independent search principles, good numerical agreement between final base-rate estimates and item-parameter values tends strongly to corroborate both (a) the latent structure postulated and (b) the numerical values thus converged upon.

7. Commencing with an item pool composed of only 20 percent moderately valid items and 80 percent "garbage" items (for separating

sexes), the methods lead to liquidation of the poor items and accurate estimates of the (p_n, p_n) psychometric parameters of the retained valid items.

8. Despite the cautionary remarks in text concerning sample size, it now appears that much smaller samples (e.g., as few as 100 patients in the taxon and 100 extrataxon) can, under favorable circumstances, yield values having 95 percent accuracy for the major latent parameters sought.

9. My conclusion, now becoming fairly firm on the growing body of Monte Carlo and real-data evidence to date, is that the Classification Society paper was, if anything, overcautious. It appears that a remarkably powerful approach exists here, deserving thorough exploration in a variety of taxonomic research problems.

REFERENCES FOR THIS CHAPTER

Alf, E., and N. M. Abrahams. Mixed group validation: A critique. *Psychological Bulletin*, 1967, *67*, 443–444.

Angyal, A., and N. Blackman. Paradoxical reactions in schizophrenia under the influence of alcohol, of hyperpnea, and CO_2 inhalation. *American Journal of Psychiatry*, 1941, *97*, 893–903.

Badia, P., A. Haber, and R. P. Runyon. *Research problems in psychology*. Reading, MA: Addison-Wesley, 1970.

Barber, B. Resistance by scientists to scientific discovery. *Science*, 1961, *134*, 596–602.

Bleuler, E. *Dementia praecox; or, the group of schizophrenias* (1911), tr. Joseph Zinkin. New York: International Universities Press, 1950.

Braatz, G. A. Preference intransitivity as an indicator of cognitive slippage in schizophrenia. *Journal of Abnormal Psychology*, 1970, *75*, 1–6.

Campbell, D. P., and D. W. Fiske. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81–105. Reprinted in E. Megargee, ed. *Research in clinical assessment*. New York: Harper and Row, 1966. Pp. 89–111. Also reprinted in W. Mehrens and R. L. Ebel, eds. *Principles of educational and psychological measurement*, Chicago: Rand McNally, 1967. Pp. 273–302. Also reprinted in D. N. Jackson and S. Messick, eds. *Problems in human assessment*. New York: McGraw-Hill, 1967. Pp. 124–132.

Carlson, E. A. *The gene: A critical history*. Philadelphia: W. B. Saunders, 1966.

Cronbach, L. J., and P. E. Meehl. Construct validity in psychological tests. *Psychological Bulletin*, 1955, *52*, 281–302. Reprinted here as Chapter 1. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-82.

Dalen, P. Causal explanations in psychiatry: A critique of some current concepts. *British Journal of Psychiatry*, 1969, *115*, 129–137.

Dawes, R. M. How clinical probability judgment may be used to validate diagnostic signs. *Journal of Clinical Psychology*, 1967, *23*, 403–410.

Dawes, R. M., and P. E. Meehl. Mixed group validation: A method for determining the validity of diagnostic signs without using criterion groups. *Psychological Bulletin*, 1966, *66*, 63–67. Reprinted here as Chapter 8.

Dorland's Illustrated Medical Dictionary. (24th ed.) Philadelphia. W. B. Saunders, 1965.

Edwards, J. H. The simulation of Mendelism. *Acta Genetica et Statistica Mediica*, 1960, *10*, 63–70.

Edwards, J. H. The genetic basis of common disease. *American Journal of Medicine*, 1963, *34*, 627–638.

Edwards, J. H. Familial predisposition in man. *British Medical Bulletin*, 1969, *25*, 58–64.

Falconer, D. S. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics*, 1965, *299*, 51–76.

Golden, R., and P. E. Meehl. Detecting latent clinical taxa, IV: Empirical study of the maximum covariance method and the normal minimum chi-square method, using three MMPI keys to identify the sexes. *Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota*. Report no. PR-73-2. Minneapolis: University of Minnesota, 1973 (in preparation). (a)

Golden. R., and P. E. Meehl. Detecting latent clinical taxa, V: A Monte Carlo study of the maximum covariance method and associated consistency tests. *Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota*. Report no. PR-73-3. Minneapolis: University of Minnesota, 1973 (in preparation). (b)

Gottesman, I. I., and J. Shields. In pursuit of the schizophrenic genotype. In S. Vandenberg, ed. *Progress in human behavior genetics*. Baltimore: Johns Hopkins Press, 1968. Pp. 67–103.

Gottesman, I. I., and J. Shields. *Schizophrenia and genetics: A twin study vantage point*. New York: Academic Press, 1972.

Heston, L. Psychiatric disorders in foster home reared children of schizophrenic mothers. *British Journal of Psychiatry*, 1966, *112*, 819–825.

Heston, L. The genetics of schizophrenia and schizoid disease. *Science*, 1970, *167*, 249–256.

Hoch, P., and P. Polatin. Pseudoneurotic forms of schizophrenia. *Psychiatric Quarterly*, 1949, *3*, 248-276.

Horst, P., ed. The prediction of personal adjustment. *Social Science Research Council Bulletin*, 1941, no. 48, 286-292.

Kuhn, T. S. *The structure of scientific revolutions*. (Rev. ed.) Chicago: Chicago University Press, 1970.

Lakatos, I. Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society*, 1968, *69*, 149–186.

Lakatos, I., and A. Musgrave, eds. *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press, 1970.

Langmuir, I. Science, common sense, and decency. *Science*, 1943, *97*, 1–7.

Linn, R. L. A note on mixed group validation. *Psychological Bulletin*, 1967, *67*, 378.

Loevinger, J. Objective tests as instruments of psychological theory. *Psychological Reports*, Monograph Supplement 9, 1957, *3*, 635–694. Reprinted in D. N. Jackson and S. Messick, eds. *Problems in human assessment*. New York: McGraw-Hill, 1967.

London, I. D. Some consequences for history and psychology of Langmuir's concept of convergence and divergence of phenomena. *Psychological Review*, 1946, *53*, 170–188.

Lykken, D. T. Statistical significance in psychological research. *Psychological Bulletin*, 1968, *70*, 151–159. Reprinted in P. Badia, A. Haber, and R. Runyon. *Research problems in psychology*. Reading, MA: Addison-Wesley, 1970. Pp. 263–277. Also reprinted in D. E. Morrison and R. Henkel, eds. *The significance test controversy*. Chicago: Aldine, 1970. Pp. 267–279.

Meehl, P. E. Some ruminations on the validation of clinical procedures. *Canadian Journal of Psychology*, 1959, *13*, 102–128. Reprinted here as Chapter 5. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. 517. (b)

Meehl, P. E. Schizotaxia, schizotypy, schizophrenia. *American Psychologist*, 1962, *17*, 827–838. Reprinted here as Chapter 7. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-516.

Meehl, P. E. *Manual for use with checklist of schizotypic signs*. Minneapolis: Psychiatric Research Unit, University of Minnesota Medical School, 1964.

Meehl, P. E. Detecting latent clinical taxa, by fallible quantitative indicators lacking an accepted criterion. *Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota*. Report no. PR-65-2. Minneapolis: University of Minnesota, May 25, 1965. (a)

Meehl, P. E. Seer over sign: The first good example. Journal of Experimental Research in Personality, 1965, *1*, 27–32. (c)

Meehl, P. E. Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 1967, *34*, 103-115. Reprinted in P. Badia, A. Haber, and R. P. Runyon. *Research problems in psychology*. Reading, MA: Addison-Wesley, 1970. (a)

Meehl, P. E. Detecting latent clinical taxa, II: A simplified procedure, some additional hitmax cut locators, a single-indicator method, and miscellaneous theorems. *Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota*. Report no. PR-68-4. Minneapolis: University of Minnesota, August 15, 1968.

Meehl, P. E. Nuisance variables and the ex post facto design. In M. Radner and S. Winokur, eds. *Minnesota studies in the philosophy of science*, IV. Minneapolis: University of Minnesota Press, 1970. Pp. 373–402. (a)

Meehl, P. E. Reactions, reflections, projections. In J. Butcher, ed. *Objective personality assessment: Changing perspectives*. New York: Academic Press, 1972. Pp. 131–189. (b)

Meehl, P. E. Specific genetic etiology, psychodynamics and therapeutic nihilism. *International Journal of Mental Health*, 1972, 1, 10-27. Reprinted here as Chapter 11. (c)

Meehl, P. E., D. T. Lykken, M. R. Burdick, and G. R. Schoener. Identifying latent clinical taxa, III: An empirical trial of the normal single-indicator method, using MMPI scale 5 to identify the sexes. *Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota*. Report no. PR-69-1. Minneapolis: University of Minnesota, January 15, 1969.

Meehl, P. E., and A. Rosen. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 1955, *52*, 194–216. Reprinted here as Chapter 2. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-514.

Morrison, D. E., and R. E. Henkel, eds. The significance test controversy. Chicago: Aldine, 1970.

Murphy, E. A. One cause? Many causes: The argument from a bimodal distribution. *Journal of Chronic Diseases*, 1964, *17*, 301–324.

Nash, L. K. *The nature of the natural sciences*. Boston: Little, Brown, 1963.

Pap, A. Reduction-sentences and open concepts. *Methodos*, 1953, *5*, 3–30.

Pap, A. *Semantics and necessary truth*. New Haven, CT: Yale University Press, 1958. (b)

Peterson, D. R. The diagnosis of subclinical schizophrenia. *Journal of Consulting Psychology*, 1954, *18*, 198–200. Reprinted in G. S. Welsh and W. G. Dahlstrom, eds. *Basic readings on the MMPI in psychology and medicine*. Minneapolis: University of Minnesota Press, 1956. Pp. 415–418.

Platt, J. I. Strong inference. *Science*, 1964, *146*, 347-353.

Popper, K. R. *The logic of scientific discovery*. New York: Basic Books, 959.

Popper, K. R. *Conjectures and refutations*. New York: Basic Books, 1962.

Rado, S. *Psychoanalysis of behavior*. New York: Grunt and Stratton, 956.

Rado, S. Theory and therapy: The theory of schizotypal organization and its application to the treatment of decompensated schizotypal behavior. In S. C. Scher and H. R. Davis, eds. *The outpatient treatment of schizophrenia*. New York: Grunt and Stratton, 1960. Pp. 87-101.

Rado, S., and G. Daniels. *Changing concepts of psychoanalytic medicine*. New York: Grunt and Stratton, 1956.

Stern, C. *Principles of human genetics*. (2nd ed.) San Francisco: W. H. Freeman, 1960.

pdf by ljy May 2000