Journal of Abnormal Psychology
1979, Vol. 88, No. 3, 217-233

#117

# Detection of the Schizoid Taxon With MMPI Indicators

Robert R. Golden and Paul E. Meehl
University of Minnesota, Minneapolis

Three different taxometric methods were used, each with a different set of MMPI schizoid indicators, in attempts to detect the schizoid taxon in a hospitalized psychiatric sample. None of the patients had been formally diagnosed schizophrenic. In each of the three analyses, a latent homogeneous class, or taxon, was detected. The three analyses were consistent in that the schizoid taxon base-rate estimates were very close (.37, .40, and .41), and the average agreement rate between methods in the classification of individuals as members of the taxon or not was as high as indicator validities would permit. The mean MMPI profile of taxon members, a 2-7-8 code type, which is usually diagnosed as schizophrenia or schizoid personality, was highly similar to the average profile of a preschizophrenic sample.

Sixteen years ago Meehl (1962) proposed a theory of schizophrenia in which he hypothesized that only a certain class of people —those who have a particular genetic constitution—have *any* liability for schizophrenia. This hypothetical class of individuals will be referred to here as the "schizoid taxon." While Meehl's theory has generated considerable interest, there has been little, if any, empirical evidence either to support or to refute it. This is probably because the taxonomic nature of the theory has resulted in methodological and statistical problems that could not be adequately handled by existing methods. The purposes of the present paper are, first, to describe three taxometric methods we have developed for testing Meehl's theory, and, second, to describe the results of our applying these methods.

The general nature of the methodological and statistical problems encountered in testing this theory was recently described (Meehl, 1973). If the specific etiology of the schizoid taxon—or "schizoidia"—is a single dominant gene, how can the probability that a person carries this gene be estimated when the only indicators available are highly fallible phenotypic ones? Currently, there is no acceptable criterion variable in the traditional sense and no diagnostically definitive touchstone, sign, symptom, or trait that we know how to measure reliably. Further, not all the correlates of schizoidia are sufficiently pathological to be called symptoms or valid enough to be called signs; hence, we will use the term *indicator*. Thought disorder or "cognitive slippage," which is viewed by Meehl (following Bleuler) as the primary indicator of schizoidia, still is not sufficient by itself for taxonomic purposes. Some clinical manifestations of cognitive slippage can be detected during intensive psychotherapy of psychiatric patients. These manifestations can be used as *inclusion* tests for schizoidia, in that their presence is a quasi-infallible indication of the presence of

---

*Editor's note.* Lewis R. Goldberg was action editor for this manuscript for which he also has my thanks. L.D.E.

This research was supported in part by National Institute of Mental Health Grant MH 24224 and grants from the Psychiatry Research Unit, the Scottish Rite Schizophrenia Research Program, and the University of Minnesota Computer Center.

The authors wish to acknowledge their indebtedness to Lewis R. Goldberg for his very active role in suggesting improvements in the present article.

Experimental versions of Fortran computer programs for each of the taxonomic methods are available from the first author upon request.

Requests for reprints should be sent to Robert R. Golden, who is now at the Division of Biostatistics, School of Public Health, Columbia University, 600 West 168th Street, New York, New York 10032.

this particular pathology. Unfortunately, these manifestations are too rare to be used successfully as *exclusion* tests. In other words, absence of these clinical manifestations does not necessarily imply absence of schizoidia. No valid psychometric tests of cognitive slippage are available—an especially serious problem for researchers who would like to study schizoidia as a hypothetical taxon that includes both schizophrenics and schizoids (who may never manifest diagnostically psychotic degrees of cognitive slippage). Therefore, we have a perfect example of a "bootstraps" problem (Cronbach & Meehl, 1955), in which we must start with a fallible set of indicators of unknown validities and hope to end up with accurate estimates of these validities on the basis of some internal statistical relations among them.

In the present paper, we first describe one of the taxometric methods we have developed. Next, we show that the method compares favorably with factor analysis and cluster analysis in solving the pseudoproblem of detecting the taxonomic variable of biological sex when sex-discriminant Minnesota Multiphasic Personality Inventory (MMPI) items are used as indicators. (We assumed that methods that could not pass this empirical trial were unlikely to be useful in detecting the schizoid taxon.) Finally, we describe the results obtained when this method was applied to MMPI data, and show how these results are consistent with the results from two other taxometric methods we have previously developed.

## The Taxometric Method

We start with an idealization (not fully obtainable in practice) that there exists a set of indicators, most of which discriminate between groups of schizoids and nonschizoids, but which are more or less independent of one another within each group. That is, we assume that ideally the indicators will intercorrelate for the mixed sample solely because the indicators discriminate between the schizoids and the nonschizoids. (The extent to which the indicators can correlate within the samples depends on the robustness of the method, which is discussed below.) Indicators that discriminate between the schizoids and the nonschizoids, and are approximately uncorrelated within the schizoids and nonschizoids, should also behave consistently with one another in various ways. For example, when each of the indicators is used to estimate a parameter, such as the taxon base rate, then all of these estimates should be similar; if they are not, we know that one or more of the indicators do not conform to the assumptions underlying the method. By the use of such internal consistency criteria, we end up with statistical tests, christened *consistency tests* by Meehl (Golden & Meehl, 1978; Meehl, 1973, 1978), which check the assumptions of the method.

A series of these consistency tests is used here in a consecutive hurdle fashion; when an indicator fails a test, it is removed and is not used in the remaining calculations. This series of tests constitutes a taxonomic search, which we call the *consistency hurdles* method. The method is iterative, in that tests are repeatedly applied to the indicators (and those failing a test are removed) until none of the remaining indicators fail any of the tests. The computational procedures and derivations are described below. More mathematical details of the derivations are provided in Golden, Tyan, and Meehl (Note 1), and in Meehl (Note 2).

### *Step 1*

Let $y_i$ $(i = 1, 2, \ldots n)$ denote a dichotomous indicator scored 1 (schizoid) or 0 (nonschizoid). Initially, the schizoid scoring direction is determined by clinical judgment or other subjective criteria. Later, the method itself checks the direction of scoring. Let $x_i$ denote the scale or "key" formed by summing the scores from all the $n - 1$ other indicators except indicator $y_i$.

$$x_i = \sum_{j=1}^{n-1} y_j \quad (j \neq i). \tag{1}$$

For each indicator $y_i$ and associated scale $x_i$, we can create the function

$$d_i(c) = a_i(c) - b_i(c), \tag{2}$$

where $c$ is a score on $x_i$, $a_i(c)$ is the mean of the $y_i$ scores for those individuals with scores above $c$ on $x_i$, and $b_i(c)$ is the corresponding mean score for those individuals below $c$.

The maximum value of $d_i(c)$ over all of the $c$ values (the possible range is 0 to $n - 1$) tells us something about the underlying latent situation: If a set of indicators are correlated mainly because each indicator is discriminative of the two taxonomic classes, then a very poor choice of $c$ should, in general, reduce the statistical tendency of $x_i$ and $y_i$ to go together. For example, if a neurologist were dealing with the clinical diagnosis of meningitis versus nonmeningitis, and if two indicators under consideration, temperature elevation and marked pain on anteroflexion of the neck, are clinical signs of meningitis, then the two signs should correlate significantly in a mixed sample of meningitic and nonmeningitic persons. But if the neurologist were so unwise as to have chosen a very low cutting point on these two indicators—such as temperature above 99° and any sign or complaint, however slight, of stiff neck or reluctance to flex the neck—then considerable numbers of patients without meningitis, but with other milder infectious conditions (including the common cold), would show one or both of the "signs." Two untoward results could be expected: first, the correlation between the two clinical indicators would be reduced, and second, the identification of the taxonomic class of interest (meningitis) would be poorer.

Monte Carlo studies of the $d_i$ function for a wide variety of artificial data samples have indicated that its maximum value over all values of $c$ must be at least .10 for the taxonomic method to give accurate results when used with indicators and taxonomies such as that of the present schizoidia study. Since, in practice, the estimate of this maximum value can be misleading as a result of sampling error (especially for samples of sizes between 100 and 500), it is best to smooth the $d_i$ curve by a method such as moving averages. Based upon these smoothed values, the first consistency test requires that max $[d_i(c)] \geq .10$. After the deletion of those $k$ indicators that fail this requirement, the $x_i$ keys are then recalculated, using the remaining set of $n - k - 1$ indicators.

*Step 2*

The $d_i$ function also has other useful properties. We will use the term *hitmax cut* for the particular (unknown) cutting score on $x_i$ (denoted by $h_i$) that maximizes the total number of correct classifications (or hits). It can be shown that the hitmax value is the cutting score associated with the point of intersection of the schizoid frequency distribution and the nonschizoid frequency distribution, but this is of no immediate consequence, since both frequency distributions are unknown, and are not directly observable. However, it has also been found by Monte Carlo studies that the value of $c$ corresponding to the maximum of the $d_i$ curve is a very good estimate of the hitmax cut, $h_i$, when the independence assumption is approximately satisfied, and when both the $x_i$ and $y_i$ indicators discriminate well between the two taxonomic classes.

Let us designate this $c$ value, which is an estimate of the true hitmax cut, as $\hat{h}_i$. And let us denote the mean of these $\hat{h}_i$ values, across all indicators, as $\bar{h}$:

$$\bar{h} = \sum_{i=1}^{n-k} \hat{h}_i/(n - k). \tag{3}$$

The second consistency test is that each of the $\hat{h}_i$ values not differ from its mean value by too much. That is, it is required that $|\hat{h}_i - \bar{h}| < e$ for each indicator $y_i$, where $e$ is another tolerance limit. Monte Carlo studies have shown that $e$ should be set equal to 15% of the range of the $x_i$ scores in the total (mixed) sample.

By now, the reader may be asking a general question about the present approach: What is the maximum amount of *inconsistency* or *disagreement* which can be tolerated by the tests? A traditional null hypothesis test to determine whether there is significant inconsistency is not desired, since we are not hopeful that literal fulfillment of the latent model will ever be possible (Meehl, 1978). In practice, the really important question is

one of robustness. We need to know how much disagreement can be tolerated in these consistency tests without incorrectly contradicting a true hypothesis that the state of nature is adequately approximated by the idealization. By "adequately approximated," we mean close enough to justify employing the method to improve our guesses as to the latent values by a worthwhile amount in relation to our initial state of ignorance. An answer to the general question of robustness is provided by the Monte Carlo and empirical studies, such as those alluded to above and those reported in more detail below.

*Step 3*

Those individuals with high scores on the $x_i$ key can be used to obtain an estimate of the proportion of schizoids that obtain a value of 1 on $y_i$; similarly, those individuals with low scores on the $x_i$ key can be used to obtain an estimate of the proportion of nonschizoids that obtain a value of 1 on $y_i$. Denote the actual values of these two proportions by $p(S_i)$ and $p(N_i)$, respectively, for the $i$th indicator. An estimate of $p(S_i)$ is given by

$$\hat{p}(S_i) = \frac{\sum y_i}{N_{\text{high}}}, \tag{4}$$

where $N_{\text{high}}$ is the number of individuals in the high $x_i$ scoring group and the numerator is the sum of the $y_i$ scores for this same group. Likewise,

$$\hat{p}(N_i) = \frac{\sum y_i}{N_{\text{low}}}, \tag{5}$$

except here we use the group of low $x_i$ scoring individuals. It is necessary to select the two subsamples of high- and low-scoring individuals so that each is large enough, in terms of sampling error, and each is taxonomically pure enough to avoid excessive bias in the estimation of $p(S_i)$ and of $p(N_i)$. Monte Carlo studies have shown that those individuals in the highest quartile of $x_i$ scores can be used for the high-scoring group, and those in the lowest quartile can be used for the low-scoring group.

We can now use these estimates to check the estimated indicator validity $\hat{p}(S_i) - \hat{p}(N_i)$. Passage of the first consistency test has nearly assured us that this quantity is at least .10. Since we have found that the method generally requires that the actual validities $p(S_i) - p(N_i)$ be at least .2, it is reasonable to require that $\hat{p}(S_i) - \hat{p}N_i) \geq .15$. The purpose of this test is to select adequately discriminant indicators, not only for accurate classification of individuals but also because this results in more accurate estimation of important parameters regarding the taxonomic classes, such as their base rates.

*Estimating the Base Rate*

After deleting indicators by the above three tests, we recalculate the $x_i$ keys from the remaining indicators and apply the tests once again. This repeated testing is continued until after all of the remaining indicators pass all of the tests. At this point, we can now estimate the schizoid base rate, $P$, the proportion of the total (mixed) sample who are schizoid. Since $P \times p(S_i) + Q \times p(N_i) = \bar{y}_i$, where $Q = 1 - P$, and $\bar{y}_i$ is the mean score on indicator $y_i$ in the total sample, we can obtain multiple estimates of $P$, by:

$$P_i = \frac{\bar{y}_i - p(N_i)}{p(S_i) - p(N_i)}. \tag{6}$$

We can take the average of these $P_i$ values as a final estimate of $P$.

If the results of these consistency tests lead to acceptance of the latent model as an adequate approximation to the state of nature, Bayes' theorem can be used to calculate the probability of each individual's membership in the schizoid taxon. Suppose, for example, that an individual has scored 1, 1, 0 on three indicators, denoted by the subscripts 1, 2, 3. Then the probability that a schizoid will have scored 1, 1, 0 is given by

$$p(1,1,0/\text{schizoid}) = p(S_1) \times p(S_2) \times (1 - p(S_3));$$

likewise, for a nonschizoid,

$$p(1,1,0/\text{nonschizoid}) = p(N_1) \times p(N_2) \times (1 - p(N_3)).$$

Table 1
*An Example Using Artificial Data*

| | Item parameters | | | | | | |
|---|---|---|---|---|---|---|---|
| | Proportion of 1 responses | | | | Consistency tests | | |
| Item ($i$) | In taxon $p(S_i)$ | Extra-taxon $p(N_i)$ | Difference $p(S_i) - p(N_i)$ | Item validity phi | Test 1 maximum $d_i(c)$ | Test 2 $\hat{h}_i - \bar{h}$ | Test 3 $\hat{p}(S_i) - \hat{p}(N_i)$ |
| 1 | .64 | .26 | .38 | .37 | .18 | −.1 | .34 |
| 2 | .84 | .48 | .36 | .38 | .22 | −.1 | .29 |
| 3 | .58 | .25 | .33 | .33 | .22 | +.8 | .34 |
| 4 | .40 | .13 | .27 | .29 | .19 | −.1 | .23 |
| 5 | .62 | .37 | .24 | .24 | .24 | −.1 | .15 |
| 6 | .90 | .71 | .19 | .24 | .19 | −.1 | .16 |
| 7 | .89 | .75 | .13 | .17 | .08[a] | — | — |
| 8 | .90 | .80 | .10 | .14 | .08[a] | — | — |
| 9 | .67 | .57 | .10 | .10 | .07[a] | — | — |
| 10 | .27 | .20 | .07 | .08 | .11 | −.1 | .07[a] |
| 11 | .62 | .56 | .06 | .06 | .11 | −.1 | .08[a] |
| 12 | .36 | .30 | .06 | .06 | .12 | −.1 | .13[a] |
| 13 | .52 | .47 | .05 | .05 | .08[a] | — | — |
| 14 | .70 | .76 | −.06 | −.06 | .02[a] | — | — |
| 15 | .74 | .81 | −.07 | −.08 | .05[a] | — | — |

*Note.* $N$ = 1,000. — = statistic not calculated since item was previously eliminated.
[a] Item eliminated as a result of failing this consistency test.

Since

$$p(\text{schizoid}/1,1,0) = \frac{Pp(1,1,0/\text{schizoid})}{Pp(1,1,0/\text{schizoid}) + Qp(1,1,0/\text{nonschizoid})},$$

the estimated probability that the individual is a member of the schizoid taxon is given by:

$$\frac{P \times p(S_1) \times p(S_2) \times q(S_3)}{P \times p(S_1) \times p(S_2) \times q(S_3) + Q \times p(N_1) \times p(N_2) \times q(N_3)} \quad (7)$$

where

$$q(S_i) = 1 - p(S_i),$$

and

$$q(N_i) = 1 - p(N_i),$$

for each value of $i$.

### *An Example Using Artificial Data*

An example will be used to illustrate the method. A random-number generator was used to obtain artificial responses, with values 0 and 1, of 1,000 individuals to each of 15 dichotomous items. Of the 1,000 individuals, 582 were actually members of one class (let us arbitrarily designate this class as the *taxon*), and the other 418 were actually members of the *extrataxon* class. For each pair of items, the (population) phi-coefficient was set equal to zero within each of the two classes. Hence, for this trial, the intraclass independence assumption, upon which the taxonomic method is based, was perfectly satisfied.

The discriminative validities for the 15 items were chosen to be similar to those likely to be encountered in actual research. For example, research with MMPI items has generally shown that items with validities over .40 are extremely rare, and few have validities over .30. Table 1 presents some characteristics of the 15 artificially produced items, ordered by their validity coefficients. These concocted validities ranged from −.08 (simulating an indicator initially scored in the wrong direction) to .37 (about as high a validity coefficient as one is likely to find with itemmetric data).

Included in Table 1 are the results from

Table 2
*Final Class Proportions and Base-Rate Estimate for the Example Using Artificial Data*

| Item | Taxon class | | | Extra-taxon class | | | Estimate of taxon base-rate |
|---|---|---|---|---|---|---|---|
| | Estimate | True value | Error | Estimate | True value | Error | |
| 2 | .83 | .84 | −.01 | .52 | .48 | .04 | .56 |
| 3 | .59 | .58 | .01 | .29 | .25 | .04 | .63 |
| 4 | .36 | .40 | −.04 | .13 | .13 | .00 | .68 |
| 5 | .61 | .62 | −.01 | .37 | .37 | .00 | .60 |
| Average | | | | | | | .62 |
| True value | | | | | | | .58 |

*Note.* $N$ = 1,000.

the first of three cycles of consistency tests. Items 7, 8, 9, 13, 14, and 15 were deleted by the first test, since their maximum $d_i(c)$ values were less than .10. The second test, that each hitmax estimate should not differ too greatly from the mean estimate, did not delete any of the nine remaining items. The third test, which requires the difference in the two class proportions to be at least .15, deleted Items 10, 11, and 12. In the second cycle, Items 1 and 6 were deleted by the second test, which concerns the hitmax interval. In the third cycle, no items were deleted. The remaining items (2, 3, 4, and 5) include four of the five that had a validity of .20 or more. From these four items, the resulting estimates of the taxonomic class proportions and the base rate, which are presented in Table 2, reflect fairly accurately the actual sample values.

The present method has built-in tests designed to tell the investigator whether there exists an underlying dichotomous taxonomy and the parameter estimates are reasonably accurate, despite assumption violations. If there is not a dichotomous taxonomy, or if the assumption of within-taxonomic-class independence is not an adequate approximation of the actual latent situation, then the errors of the parameter estimates may be too large, and may possibly mislead the investigator. The method has been studied by the Monte Carlo technique with regard to robustness (sensitivity to assumption violation) and spuriousness (detecting a taxonomic class when none exists). Such Monte Carlo trials indicate that, while the method is not highly robust, it is not likely to produce a spurious detection. To test the tendency to give spurious results, the method was used to analyze various samples of nontaxonomic but continuous data generated according to a latent trait model (Lord & Novick, 1968). For a wide variety of samples of such artificial data, all of the items were always deleted by the method.

### *Empirical Trials of the Method: Differentiating the Sexes*

An empirical test in which MMPI items were used to identify the sexes suggests how well the method works with real data. Samples of 430 men and 720 women were used to compare the sexes on all 550 MMPI items. The items were keyed 1 for the female direction and 0 for the male direction, the keying direction being determined empirically using the present samples. Fifteen items discriminated the males from the females with a difference in item proportions of .30 or above, 12 items did so with a difference of .20 to .30, and 15 items did so with a difference of .10 to .20. In addition to these 42 discriminating items, 33 nondiscriminating items (with differences between −.10 and .10) were randomly selected from the remaining 508 items, in order to simulate an actual research situation, such as detecting the schizoid taxon, where many invalid items may be tried. The combined set of 75 items was then used to attempt to detect the taxonomy of biological sex.

Before presenting the results obtained by

applying the present taxonomic method to these data, let us consider the data obtained by using two other popular item-clustering methods. The hierarchical-cluster method developed by Ward (1963) was used with the 75 items and a mixed sample of 100 males and 100 females; it did poorly in the identification of the sexes at all stages of the iteration procedure. For the partition where there were two clusters, the proportion correctly classified from cluster membership was only .64. Similarly, inverted factor analysis developed by Stephenson (1938) failed to produce factors that were significantly more highly loaded for members of one sex than the other.

When the present method was applied to these data, 28 of the 75 items had maximum $d_i(c)$ values greater than .10 (Consistency Test 1) on the first cycle of testing. Further testing of these 28 items eliminated all but seven. Table 3 presents the results based on these seven items. Note that the male and female class proportions were estimated fairly accurately (usually within .10) and the base rate was estimated to be .67, which is close to the sample value of .63. There is obviously room for improvement in the method, since it retained less than one fourth of the valid items. However, the method did provide estimates of the base rate and the class proportions which, when used with Bayes' theorem, yielded a correct classification rate of nearly 90%.

The results presented in Table 3 are based on 1,150 subjects. However, since only 211 subjects were available for the schizoidia analyses (to be presented below), it was important to check the accuracy of the method with a sample of comparable size. Consequently, the analyses were carried out once again, this time using 100 males and 100 females. The class proportions and base-rate estimates for this sample of 200 subjects are presented in Table 4. As one can see, even with a sample of this size, the item class proportions were estimated quite accurately, as was the base rate (.54 versus .50).

## The Schizoidia Study

According to Meehl (1962), an individual who has a genetic predisposition to become schizophrenic also has four cardinal personality traits: cognitive slippage, social aversiveness, anhedonia, and ambivalence. The only direct phenotypic consequence produced by the hypothesized genetic aberration is an *integrative neural defect,* which is due to some parameter of single cell function. Whether or not this defect is manifested in more molar functions of the central nervous system depends on the organization of the mutual feedback controls and on the stochastic parameters of the reinforcement regime. Only this neural integrative defect, christened *schizotaxia* by Meehl, can properly be spoken of as inherited. The effects of

Table 3
*Final Class Proportions and Base-Rate Estimate for the 75-Item Male–Female Trial*

| | Males | | | Females | | | |
|---|---|---|---|---|---|---|---|
| Item | Estimate | True value | Error | Estimate | True value | Error | Estimate of female base rate |
| 1 | .08 | .06 | .02 | .48 | .69 | −.21 | .74 |
| 2 | .24 | .31 | −.07 | .67 | .74 | −.07 | .68 |
| 3 | .14 | .15 | −.01 | .44 | .55 | −.11 | .71 |
| 4 | .04 | .11 | −.07 | .53 | .54 | −.01 | .58 |
| 5 | .48 | .48 | .00 | .63 | .77 | −.14 | .68 |
| 6 | .08 | .27 | −.19 | .33 | .34 | −.01 | .66 |
| 7 | .36 | .55 | −.19 | .74 | .69 | .05 | .69 |
| Average | | | | | | | .67 |
| True value | | | | | | | .63 |

*Note.* $N$ = 1,150.

Table 4
*Final Class Proportions and Base-Rate Estimate for 75-Item Male–Female Trial*

| Item | Males | | | Females | | | Estimate of female base rate |
|---|---|---|---|---|---|---|---|
| | Estimate | True value | Error | Estimate | True value | Error | |
| 1 | .05 | .08 | −.03 | .66 | .64 | .02 | .53 |
| 2 | .09 | .10 | −.01 | .37 | .30 | .07 | .57 |
| 3 | .41 | .50 | −.09 | .87 | .90 | −.03 | .55 |
| 4 | .42 | .44 | −.02 | .84 | .74 | .10 | .46 |
| 5 | .46 | .40 | .06 | .91 | .84 | .07 | .48 |
| 6 | .25 | .36 | −.11 | .87 | .78 | .09 | .54 |
| 7 | .05 | .14 | −.09 | .54 | .52 | .02 | .56 |
| 8 | .08 | .08 | .00 | .47 | .38 | .09 | .55 |
| 9 | .30 | .36 | −.06 | .80 | .84 | −.04 | .57 |
| Average | | | | | | | .54 |
| True value | | | | | | | .50 |

*Note.* $N = 200$.

a history of social learning upon schizotaxic individuals results in a personality organization which Meehl called the *schizotype*. Meehl believes that the four core behavior traits are not innate in schizotaxic individuals, but rather that they *learn* these traits, regardless of their particular social reinforcement histories (which can range from best to worst). If the learning history in interpersonal behavior is favorable, and the schizotaxic person also is fortunate enough to have inherited physical vigor and high thresholds for anxiety and stress, then he or she should remain a well-compensated schizotype, never manifesting clinical symptoms of schizophrenia. A compensated schizotype is analogous to a gout-prone male, who has a genetically determined elevated blood – uric acid titer but never develops clinical gout.

Meehl (1962) further hypothesized that the statistical relation between schizotaxia, schizotypy, and schizophrenia is that of class inclusion. In other words, *all* schizotaxics become schizotypic in personality organization, regardless of the quality of their particular social learning regimes, but most will remain compensated. A minority, disadvantaged by additional (largely polygenically determined) constitutional weaknesses and by a history of poor social learning influenced by schizophrenogenic mothers (most of whom are themselves schizotypes), become clinical schizophrenics. Meehl postulated that schizotaxia is etiologically specific; it is a necessary condition for the development of schizophrenia. Nonschizotaxic individuals, whatever their other genetic makeup and learning histories, will at worst develop character disorders, psychoneuroses, or other psychoses (e.g., manic depression). Lacking the integrative neural defect, they will not become schizotypes, and therefore can never manifest its decompensated form, schizophrenia.

We conjecture only that a taxometrically identified class will turn out to be closely related to, but not identical with, the schizotype. Other reasons for the use of an additional construct are that the present MMPI data cannot provide a test for the etiology of any taxon we might detect, and additionally we do not have MMPI measures of the four cardinal traits. However, if our strongest conjectures turn out to be substantially correct, as confirmed by evidence from further studies, the schizoid taxon will correspond to the schizotype.

We assume that an item should discriminate between schizophrenics and normals if it does so between schizoids and nonschizoids. This was the first requirement used in selecting candidate items. Of the 550 MMPI items, 53 were found to discriminate between 96 diagnosed schizophrenics and the MMPI normal sample by a difference in the class proportions of .20 or more.

The next requirement was that the item

not be highly correlated with decompensation-related variables, such as severity of the illness. The purpose of this requirement was to prevent the selection of items that were highly correlated with each other, within the schizoids, through underlying decompensation-related variables. Moreover, items were desired that did not discriminate among other diagnostic classes (e.g., psychotics vs. neurotics). In the present study, we were able to require only that an item not discriminate highly among diagnosed subtypes of schizophrenia, or among other psychoses for which we had samples of sufficient size. This last requirement was failed by 20 of the 53 previously selected items, leaving 33 items for further analysis.

The sample used for the attempted detection of the schizoid taxon consisted of 211 male inpatients at the University of Minnesota Hospital who had diagnoses of neurosis, personality disorder, or transient situational disorder. No schizophrenics or patients with other psychotic diagnoses were included in this sample, in order to reduce the possibility of detecting taxonomic classes other than the one of interest.

When the present method was applied to this sample, using the 33 selected MMPI items, it deleted all but 7 items. These items, along with their parameter estimates, are listed in Table 5. The taxon base rate was estimated to be .37. Using these parameters and Bayes's theorem, individuals were classified as either schizoid or nonschizoid. The Bayes probabilities tended to be either close to zero or close to one, a result that previous Monte Carlo analyses have shown to be indicative of a real taxonomy.

The average correlation between all pairs of the 7 items in the mixed sample was .15. (For comparison, if all items are uncorrelated within the taxonomic classes, if the base rates are each .50, and if all differences between the class proportions are .30, then the average item intercorrelation for the mixed sample will be .16.) The average item intercorrelation was −.08 among those classified as schizoid, and −.01 among those classified as nonschizoid. The method evidently selected 7 items which satisfied the independence assumption.

The two subsamples resulting from classifying each patient as schizoid or nonschizoid differed by .20 or more in their endorsement proportions on 113 of the 550 MMPI items. Thirty of these 113 items discriminated between the two subsamples by .30 or more. Strangely, none of these 30 items were included in the original 53 items that discriminated between the schizophrenics and the normals. When these 30 items were analyzed by the present method, 8 items were retained, and the estimated schizoid base rate was .38, very close to the previous estimate of .37. When the two sets of item

Table 5
*Seven MMPI Items Selected by the Taxonomic Method*

| MMPI booklet no. | Schizoid direction | Schizoid proportion | Nonschizoid proportion | Difference | Item |
|---|---|---|---|---|---|
| 61 | T | .75 | .39 | .36 | I have not lived the right kind of life. |
| 239 | T | .65 | .31 | .34 | I have been disappointed in love. |
| 20 | F | .52 | .19 | .33 | My sex life is satisfactory. |
| 317 | T | .65 | .34 | .31 | I am more sensitive than most other people. |
| 284 | T | .61 | .32 | .29 | I am sure I am being talked about. |
| 501 | F | .42 | .22 | .20 | I usually work things out for myself rather than get someone to show me how. |
| 207 | F | .37 | .21 | .16 | I enjoy many different kinds of play and recreation. |

*Note.* T = true; F = false. MMPI = Minnesota Multiphasic Personality Inventory.

indicators were used separately to classify individuals as *in* or *out* of the schizoid taxon, they agreed for 70% of the sample. While an agreement rate of .70 may seem low, it is nearly as high as could have been expected. If one method correctly classifies a proportion, $p_1$, of the total mixed sample, while a second method does so with a proportion $p_2$, and if the two methods make *independent* errors of classification, then the proportion of classifications for which the two methods agree is $p_1p_2 + (1 - p_1) \times (1 - p_2)$. If $p_1 = p_2 = .80$, a rather high degree of accuracy, then the agreement rate would be .68.

An estimate of the correct classification rate, $p$, for a single set of indicators can be obtained by a method developed in Golden and Meehl (Note 3). For each individual, we used Bayes's theorem to calculate the probability of being schizoid and that of being nonschizoid and selected the larger of these two values. It can be shown that an estimate of the overall correct classification is the average of the larger of these two probability values across all individuals. The estimated correct classification rate for each of the two sets of items was found to be about .85, which is quite high. The estimated agreement rate between the two classifications should then be about $.85 \times .85 + .15 \times .15 = .75$, which is close to the observed values of .70. Hence, the concordance seems in satisfactory agreement with what the separate classification estimates would predict.

The small number of items retained by the method deserves some discussion. The main reason for this result is evidently the strong assumption of independence. Another likely reason is the choice of tests and tolerance limits to minimize the possibility of spurious results. Although classification with a small number of items may seem hazardous, as few as five items that are quasi-independent within each taxonomic class, and moderately discriminative between classes, provide a quite accurate classification of individuals. For example, it can be shown algebraically that, given equal class base rates, five items with validities of .20 that are mutually independent within each class will produce a misclassification rate via Bayes' theorem of only .20. Indeed, for five items with validities of .30, the misclassification rate will be just .10, while for five items with validities of .40, it will be a mere .05. That is, a few powerful but independent items may discriminate better than many weak items. This should come as no surprise, despite the tendency in traditional psychometrics to use large numbers of weak items. Physicians have found it more useful in the diagnosis of medical problems to concentrate on a few strong signs of the various nosological entities, rather than to use summations of many feeble ones.

## Other Taxometric Analyses

In addition, the MMPI item data were analyzed by two other methods which had been previously developed by the present authors. Each of these taxonomic methods was used to attempt to detect the schizoid taxon, but with different sets of MMPI items than those used above. The Maxcov (Meehl, 1973; Meehl, Notes 2 and 4; Golden & Meehl, Notes 3 and 5) and Normal taxonomic methods (Hasselblad, 1968; Golden, Tyan, & Meehl, Note 6) each require the use of continuous, as opposed to dichotomous, items as indicators and employ consistency tests after all latent parameters have been estimated. Both methods require the use of indicators that adequately discriminate between the two taxonomic classes, and that are not highly correlated within each of the classes. Both methods provide estimates of the taxon base rate and other class parameters.

The Maxcov method is based on computing the covariance between two indicators in the total (mixed) sample for each small interval of values on a third, or controlling, indicator. Under certain assumptions, the covariance obtains a maximum value (except for sampling error) at the hitmax cut on the controlling indicator. It follows that the interval of the controlling indicator associated with this maximum covariance contains approximately a half-and-half mixture of the two taxonomic classes, permitting one to estimate the latent marginal frequency distributions of each class. By letting each of the indicators play the role of the controlling

indicator, the latent distributions on each of the indicators are estimated.

Three scales were constructed from the 113 items that discriminated between the detected taxonomic classes of schizoids and nonschizoids by a difference of .20 or more. Factor analysis with varimax rotation of the 113 items in the total mixed sample ($N = 211$) suggested three factors; the 20 highest loading items for each factor were selected to form the three scales. When these three keys were used with the Maxcov method, the taxon detected had a base-rate estimate of .40, which agrees well with that of the first method. The Maxcov method also includes consistency tests that help one avoid being misled by inaccurate parameter estimates, and Monte Carlo study indicates that they work quite well (Golden & Meehl, Note 5). Passing these tests provides additional support for the theory that the detected taxon is not spurious.

The third taxonomic method, which we call the Normal method, can be used with a single indicator. In this method it is assumed that the frequency distribution in the total (mixed) sample is a mixture of two latent overlapping normal frequency distributions, one for each taxonomic class. The means, variances, and base rates of the latent distributions can be estimated by the maximum-likelihood method (Hasselblad, 1968). One way to test this method is to compare the observed mixed frequency distribution to that estimated from the latent parameters by use of the chi-square test. A low value of chi-square corresponds to a good fit of the model.

Factor analysis of the 13 standard MMPI scale scores in the total sample ($n = 211$) produced a varimax factor that accounted for 41% of the common variance and that correlated highly with the Psychasthenia (.69), Schizophrenia (.53), Depression (.61), and Social Introversion (.79) scales; all other loadings were below .30, except for $K$ (−.43). The items in these four scales were combined to make a long scale which was used as the single indicator. The schizoid taxon base rate was estimated by the Normal method to be .41, again in excellent agreement with previous estimates. The chi-square value of 2.3 was nonsignificant and was even below the value expected if the assumptions of the method were perfectly satisfied. The difference between the two taxonomic class means on this indicator was estimated to be about two class standard deviations. Since this is about the same degree of separation obtained for the sexes in the previous analysis using MMPI items, we have additional evidence that the schizoid taxonomy appears to be detectable with MMPI items as indicators.

All three of the taxonomic methods classified individuals as either *in* or *out* of the schizoid taxon, with agreement rates between pairs of methods very near to what would be expected from the estimates of their misclassification rates. This result obtained even though the three methods are based on quite different assumptions, and were used with different MMPI indicators. Also, the three base-rate estimates were remarkably close to each other and to a personal clinical estimate of .40 to .45 by Meehl; the latter was a prerecorded "impressionistic guess" based on some 30 years of outpatient private practice and the use of the Checklist of Schizotypal Signs (Meehl, Note 7).

## The Nature of the Schizoid Taxon

Each individual whose probability of being a member of the schizoid taxon exceeded .5 was classified as *probable schizoid* and others as *probable* nonschizoid, thereby forming two subsamples that could be compared on other variables. For this purpose we had three taxonomic results and we selected the set of seven items resulting from the first application of the consistency-hurdles method. It is now evident that classification on the basis of all three of the methods may have significantly improved the reliability and validity of this classification.

The two subsamples had approximately the same mean age (36.0 and 35.5 yr), and approximately the same distribution of psychiatric diagnoses. Prior to the present study, 124 MMPI items had been selected by Meehl, based on his clinical experience, as having face validity for discriminating schizoids. A comparison of those classified as

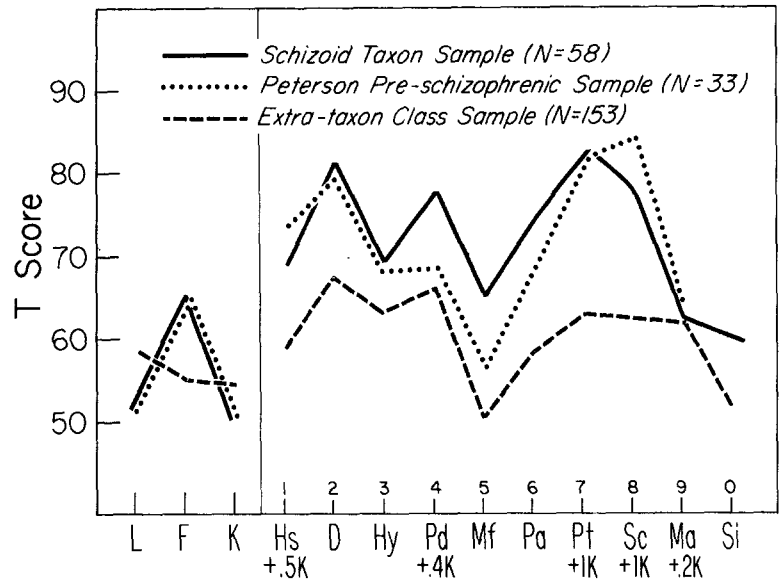


*Figure 1.* Mean MMPI profiles for taxometrically identified schizoids and nonschizoids and for clinically identified preschizophrenics. (Letters refer to MMPI scales.)

schizoids with those classified as nonschizoids on these 124 items revealed that 108 items had a difference in the anticipated direction, the average difference being about .20.

When the two subsamples were compared on the standard MMPI scales, those classified as schizoid taxon members had higher mean scores on all ten scales. The largest difference was on Psychasthenia (about 1.3 within-class *SD*) followed by Hysteria, Psychopathic Deviance, Masculinity–Femininity, Paranoia, Schizophrenia, and Social Introversion (about 1 *SD* for each), and then by Depression, Hysteria, and Hypomania (about .5 *SD* for each).

The mean MMPI profile for the individuals classified as members of the schizoid taxon, presented in Figure 1, was nearly identical to the 2–7–8 code type; using either the rules of Marks and Seeman (1963) or those of Gilberstadt and Duker (1965), the mean profile met 14 of the 18 requirements and barely failed the remaining 4, even though only a small percentage of the *individual* profiles were of the 2–7–8 type. This finding is consistent with the findings of Goldberg (1972), who argued that the mean profile of a homogeneous group represents the underlying pathology common to the individuals better than do the individual profiles, because of the inherent unreliability of the latter. A study by Briggs, Taylor, and Tellegen (Note 8) showed that, for a sample of 2,875 patients from the same population as those in the present sample, less than 5% had 2–7–8 profiles according to the Marks and Seeman rules, even when one rule violation was allowed. It follows that if the base rate of schizoidia is about .40 among nonschizophrenic psychiatric patients, the code type 2–7–8 cannot be relied upon for such a diagnosis.

The mean MMPI profile for those individuals classified as *not* members of the schizoid taxon was considerably lower, and not similar to any standard code type. This subsample is presumably quite heterogeneous, as one would expect.

The most impressive evidence of construct validity was that the mean MMPI profile of the schizoid taxon was very similar to that of a sample of preschizophrenics studied by Peterson (1963). A search of Lanyon's (1968) *Handbook of Group MMPI Profiles* revealed few nonschizoid diagnostic groups that also have the 2–7–8 mean profile pattern exhibited by both the present schizoid taxon and the Peterson preschizophrenic sample.

Finally, for the total sample ($N = 211$), the 7-item schizoidia scale correlated .56 with the Psychasthenia, .53 with the Schizophrenia, .41 with the Social Introversion, and .42 with the Depression scales. The scale correlated .81 with the long key which was used as the single indicator for the Normal method. Since this latter value is very near the limit imposed by the imperfect reliabilities of the two measures, it would appear that in practice the indicator consisting of the sum of four standard scale scores (Depression, Psychasthenia, Schizophrenia, and Social Introversion) will work as well as any other MMPI index. This result allows, first, for attempts at replication of the present findings in large samples where only scale scores have been recorded and, second, for using the more reliable sum of four scales for the diagnosis of individuals.

Despite the reassuring convergence of evidence, we are uneasy. Since this unease is shared with several readers and auditors of the paper, its causes should be identified and, if possible, dispelled. Perhaps the best lead to understanding one's cognitive discontent is a question put by several critics on hearing or reading drafts of the paper: "Now really, do you fellows seriously propose to diagnose individual patients as belonging in the schizoid taxon, or falling outside it, by examining their responses to these seven MMPI items?" We find ourselves reluctant to answer that question with a clear affirmative; yet, it seems that we ought, in consistency, to be willing to do so. We have, after all, calculated individual inverse probabilities from Bayes's formula and have estimated the true (construct-valid) latent hit-rate of such classifications. We have argued that the striking U-shaped distribution of these probabilities over individuals in our sample strongly corroborates the taxonic hypothesis and the associated hypotheses assigning high validity to the indicators employed. If this reasoning is substantially correct, why should one experience hesitation in advocating the diagnosis of individuals (found in a psychiatric population) as schizotypal or not, especially in cases where a patient's 7-item response pattern puts him or her in one of the two cusps at the ends of the distribution of Bayesian inverse probabilities?

We discern four related worries that give rise to the critics' skepticism and our own ambivalence about employing the items for individual diagnosis. While they are interrelated in important respects, the four are not equivalent, so they have a cumulative impact. The first objection has been voiced as a question about how one can possibly view such a small set of only seven items, and verbal-structured inventory items at that, as being anything like an adequate sample of the schizotypal item domain. Our answer to that, which is not a foolish question (but, we think, based on an unacceptable premise in this context), is that we are not viewing this small set of seven items as one tends to view items in classical psychometrics—especially the psychometrics of achievement and ability tests—as samples of a domain. In terms of the distinction made by Goodenough as relied on by Cronbach and Meehl (1955, p. 281), these seven items are to be viewed as *signs* rather than *samples*. This is not the place to develop in detail the conceptual and mathematical differences between signs and samples, nor do we argue that the distinction is always clear-cut or philosophically fundamental. When the internist or neurologist uses a relatively small set of high-weight indicators in arriving at a medical diagnosis, he or she does not think of them as "statistical samples of a symptom domain" in anything like the way we have customarily thought in classical psychometrics. One of the more obvious respects in which he or she proceeds differently is by using symptoms that are not highly correlated. A psychologist, however, thinking traditionally would not be happy with a set of items on a spelling test, or even a test of social introversion, that did not have a high internal consistency, as measured by some appropriate statistic such as Cronbach's Alpha. If each of our seven items truly possesses the moderate to high discriminating power reported in the table for this sample, and is negligibly correlated within the taxon and the extrataxon class, it is an inescapable consequence—a matter of algebra

—that the Bayesian inverse probabilities computed by relying on some of the 128 possible patterns will be close enough to zero or to one that we can assign class membership to individual patients with considerable confidence. In so doing, we think of the seven psychometric indicators not as samples from a statistically homogeneous *domain of items* but as signs of an *inner condition*. These signs function analogously to a relatively small number of relatively high-weight symptoms in organic medicine.

Part of the worry over this small number of items seems to be caused by the item content, which one does not read as being very "crazy" and hence not likely to be powerful for the detection of schizotypy. The first reassurance to be given on that subject is that we are not identifying florid schizophrenics, who were deliberately excluded along with the affective and brain syndrome psychoses from the mixed psychiatric sample in the present study. No one familiar with the literature on pseudoneurotic schizophrenia, subclinical schizophrenia, borderline states, and the like, should be surprised to find subtle indicators employable for such a purpose. Nor should anyone familiar with the MMPI literature and the theory of this test be distressed even if, after the fact, one can't completely satisfy oneself as to the alleged "schizospecificity" of an item. It is, for instance, quite possible that in the natural history of the disease, there are not only declines in item validities, or rises from invalidity to validity, as a patient passes from the compensated or semicompensated schizotypal status to a disintegrated schizotypal status ("psychosis" in Rado's language); it is conceivable on general theoretical grounds and clinical experience that a few items might even reverse themselves.

We are not as happy with the content of these seven items as we might have hoped to be. But we are not astounded by this content either. We suggest that clinicians who have done appreciable amounts of intensive psychotherapy with pseudoneurotic schizophrenias and schizoid personalities will find the content at least acceptable, in that it is not *dis*harmonious with the phenomenology and theory of the disorder. If one reflects on one's understanding of the phenomenology and psychodynamics of schizophrenia insofar as it is presently understood—quite apart from genetic and other etiological considerations—we think these will be found acceptable items, granted that there are other items in the MMPI pool that might strike one as equally acceptable. It is not obvious that there are items in the MMPI pool that are more acceptable. We are looking for items that discriminate the decompensated schizotype who shows up with a formal diagnosis of something other than schizophrenia or an affective psychosis. Therefore we do not expect any items to "work" in that discrimination by expressing the core symptoms of a clinically recognizable florid schizophrenia, such as clearly inappropriate affect, autism, markedly disturbed associations, schizophasic speech, or expressing Bleuler's accessory symptoms of delusions, hallucinations, catatonic symptoms, and the like. Items reflecting disappointment in life and love (the first three in the list), items reflecting interpersonal aversiveness (the next two in the list), and an item reflecting restriction of interest and perhaps, indirectly, anhedonia (the last one about play and recreation) fit easily into our picture of the schizotype. The remaining item (preferring not to work things out for oneself) we are at a loss to explain without engaging in more after-the-fact ad hockery than seems desirable.

A third source of discomfort is that the taxon is intended to refer to the schizotypal genome, and there is something inherently implausible about a verbal item having a high diagnostic power for an inferred genome. As we have found ourselves saying, "These MMPI item responses are too many steps removed in the causal chain from the DNA." But we conclude that this objection is less serious than it appears, provided the other objections are answerable. It does not matter how many causal steps intervene between a cistron and a phenotypic indicator, so long as the probabilities are sizable *and numerically estimatable by multiple methods*. One has a tendency, understandable but we think unsound, to emphasize the long and complex

"causal history" that goes into the acquisition of the verbal behavior needed to respond immediately to an MMPI item. But this historical complexity in acquiring language descriptive of one's inner events, or commenting on one's interpersonal relationships, should not be conflated with the length and complexity of the *presently functioning* causal linkage between the DNA and the verbal report as given by an adult who has acquired language. The causal linkage between the schizogene and, say, a deficit in hedonic capacity might be relatively short and direct (e.g., biochemistry of dopamine). Then all that we really require is that a person has previously learned to use appropriate self-descriptive vocabulary. That verbal skill being present, a patient's response to an MMPI item characterizing something about his inner state might be a very direct consequence of whatever brain abnormality is the endophenotype for schizoidia.

A medical analogy: One might similarly doubt the relevance of a patient's verbal report of excruciating pain in the back and scrotum as a powerful indicator of renal calculus, arguing that the acquisition of language descriptive of one's inner bodily conditions, the characterization of pains as lancinating or throbbing or dull, the grading of intensity, the adequacy of localization, and so forth, are too complex and probabilistic for clinicians to rely on language characterizing a referred pain to the testicle, when the lesion is in the ureter, and the like. But surely such doubts would be misplaced. We repeat that the question is not whether a complex causal history was involved, but whether, *once that causal history of language acquisition has in fact occurred,* the indicator weight of a subjective report or a self-observation on one's social behavior, as bearing on the inner events or structures, is strong or weak. Another trite, but illuminating, example would be a personality inventory item that says, "I am a woman," which presumably (absent a very unusual psychotic delusion or somebody deliberately fooling the examiner) would have a quasi-perfect indicator validity for the XX genotype. So we believe that the tenuous inference from a few bits of verbal behavior to an alleged schizotypal make up, and hence to an inferred schizogene, is not as weighty an objection as it first appears.

A fourth and last source of worry is the small sample size, now not of items, but of patients. A sample size of around 200 with a taxon base rate of around .4 means that we are estimating latent parameters such as the valid positive rate ($p_s$) and the false positive rate ($p_n$) relying on a taxon frequency of only around 100 cases. It is well known that the sampling instability of proportions is discouragingly high. For such a sample size, the standard error of a proportion is around .05, and it seems dangerous to use proportions with that degree of sampling instability by putting them in Bayes's formula and multiplying them out. The authors disagree between themselves as to the seriousness of this problem. A number of Monte Carlo runs and analytic considerations not reported here provide some reassurance on this score. However, it must be stated clearly that our taxometric methods were not intended to be used with samples of this small size. In general, our methodological emphasis on point estimates and consistency tests (in preference to the usual significance-testing approaches in the social sciences) tends naturally to lay heavy stress on sample size. A significance test, although it may not have adequate power and hence may lead to errors of Type II if interpreted incautiously, will not by its very mathematical nature lead to a serious number of Type I errors for a given significance level. Consequently, there is a sense in which small samples are acceptable in that tradition. But, there is no method known to statistics, and never can be one, that will give high sampling stability to a number, like a proportion, estimated from a small sample. If one is interested in the convergence of point predictions from various methods of making the estimates, one should, theoretically, insist on sufficiently large samples so that sampling error as such will not be a major contributor to a failure of consistency.

On the other hand, it can plausibly be argued that while many MMPI items initially considered by us were, in fact, good

items, they were eliminated because of their failure to satisfy the consistency-hurdles criteria employed. So, even though an individual item that survived the procedure could have suffered from a sizable sampling deviation in the valid or false positive rate, the items that did survive all hurdles would be those that did not err grossly. In rejoinder to this rebuttal, it might be urged that multiple consistency hurdles could be capitalizing on sampling errors in subtle configural ways not yet understood. The authors have agreed to disagree about the weightiness of this small sample objection, the second author offering a dissenting opinion that if there is anything seriously wrong with the present article, it lies in our willingness to publish it on the basis of $N = 211$. If we or others fail to replicate the results in the sense of not finding exactly the same set of strong items (or an equally good set), we think this would show that we erred in our reliance upon a sample of this size for these purposes, but that such an adverse result will not tend to undermine the basic reasoning of the method. As was pointed out at the beginning, the purpose of the present paper is primarily conceptual and methodological, rather than substantive. We intend to repeat the whole procedure on a sample composed of several thousand cases, in which it can be assumed that the latent parameters $P$, $p_s$, $p_n$, and the correlations of items pair-wise will not suffer from instability due to sample size.

## Conclusions

It is likely that there exist neurological and physiological variables that will prove to be much more powerful indicators of schizoidia than MMPI responses. However, MMPI responses, even though far removed causally from any genetic etiological source, may still suffice, in samples of a thousand or more, to provide a convincing test of the existence or nonexistence of the schizoid taxon. The results of the present preliminary trial, even though the sample size is quite small, indicate that such a taxon is likely to exist. These results are sufficiently encouraging to justify an attempted replication of the present study with a much larger sample. The major purpose of this paper has been to describe a promising new kind of methodology. The substantive results presented here are not regarded as confirmation of a theory, but are offered in the "context of discovery" as an invitation to attempt to replicate and extend them.

## Reference Notes

1. Golden, R. R., Tyan, S. H., & Meehl, P. E. *Detecting latent clinical taxa, VI: Analytical development and empirical trials of the consistency hurdles theory* (Rep. PR–74–4). Minneapolis: University of Minnesota, Deports from the Research Laboratories of the Department of Psychiatry, 1974.
2. Meehl, P. E. *Detecting latent clinical taxa, II: A simplified procedure, some additional hitmax cut locators, a single-indicator method, and miscellaneous theorems* (Rep. PR–68–4). Minneapolis: University of Minnesota, Reports from the Research Laboratories of the Department of Psychiatry, 1968.
3. Golden, R. R., & Meehl, P. E. *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* (Rep. PR–73–3). Minneapolis: University of Minnesota, Reports from the Research Laboratories of the Department of Psychiatry, 1973.
4. Meehl, P. E. *Detecting latent clinical taxa by fallible quantitative indicators lacking an accepted criterion* (Rep. PR–65–2). Minneapolis: University of Minnesota, Reports from the Research Laboratories of the Department of Psychiatry, 1965.
5. Golden, R. R., & Meehl, P. E. *Detecting latent clinical taxa, V: A Monte Carlo study of the maximum covariance method and associated consistency tests* (Rep. PR–73–4). Minneapolis: University of Minnesota, Reports from the Researh Laboratories of the Department of Psyhiatry, 1973.
6. Golden, R. R., Tyan, S. H., & Meehl, P. E. *Detecting latent clinical taxa, VII: Analytical development and empirical and artificial data trials of the multi-indicator, multi-taxonomic class maximum likelihood normal theory* (Rep. PR–74–6). Minneapolis: University of Minnesota, Reports from the Research Laboratories of the Department of Psychiatry, 1974.
7. Meehl, P. E. *Manual for use with checklist of schizotypic signs.* Minneapolis: University of Minnesota Medical School, Psychiatry Research Unit, 1964.
8. Briggs, P. F., Taylor, M., & Tellegen, A. *A study of the Marks and Seeman MMPI profile types as*

*applied to a sample of 1,875 psychiatric patients* (Rep. PR-66-5). Minneapolis: University of Minnesota, Reports from the Research Laboratories of the Department of Psychiatry, 1966.

## References


Cronbach, L. J., & Meehl, P. E. Construct validity in psychological tests. *Psychological Bulletin,* 1955, *52,* 281–302.

Gilberstadt, H., & Duker, J. *A handbook for clinical and actuarial MMPI interpretation.* Philadelphia: Sanders, 1965.

Goldberg, L. R. Man versus mean: The exploitation of group profiles for the construction of diagnostic classification systems. *Journal of Abnormal Psychology,* 1972, *79,* 121–131.

Golden, R., & Meehl, P. E. Testing a dominant gene theory without an accepted criterion variable. *Annals of Human Genetics* (London), 1978, *41,* 507–514.

Hasselblad, V. Estimation of parameters for a mixture of normal distributions. *Technometrics,* 1968, *8,* 431–444.

Lanyon, R. I. *A handbook of MMPI group profiles.* Minneapolis: University of Minnesota Press, 1968.

Lord, F. M., & Novick, M. R. *Statistical theories of mental test scores.* Reading, Mass.: Addison-Wesley, 1968.

Marks, P. O., & Seeman, W. *Actuarial description of abnormal personality.* Baltimore: Williams & Wilkins, 1963.

Meehl, P. E. Schizotaxia, schizotypy, schizophrenia. *American Psychologist,* 1962, *17,* 827–838.

Meehl, P. E. MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. In P. E. Meehl, *Psychodiagnosis: Selected papers.* Minneapolis: University of Minnesota Press, 1973.

Meehl, P. E. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology,* 1978, *46,* 806–834.

Peterson, D. R. The diagnosis of subclinical schizophrenia. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: University of Minnesota Press, 1963.

Stephenson, W. The inverted factor technique. *British Journal of Psychology,* 1938, *26,* 344–361.

Ward, J. H., Jr. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association,* 1963, *58,* 236–244.


Received December 11, 1978 ■