ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 74, No. 1, April, pp. 53–61, 1998
ARTICLE NO. OB982752

Nonlinear Models in Decision Making: The Diagnosis of Psychosis versus Neurosis from the MMPI

Yoav Ganzach

Tel Aviv University, Tel Aviv, Israel

Previous studies have found no nonlinear relationships between the MMPI scales and the diagnosis of psychosis versus neurosis (the criterion). However, these studies also found evidence for nonlinear accuracy in MMPI-based judgments of the likelihood of psychosis. The current study identifies the rules underlying this nonlinear accuracy and demonstrates their validity in predicting the criterion. © 1998 Academic Press

Are there useful nonlinear rules to predict decision criteria from relevant predictors? That is, are there useful rules in which the predicted criterion is *not* a linear combination of the predictors? Much of the research on this question was conducted using the MMPI scales as predictors and the diagnosis of psychosis versus neurosis as a criterion (e.g., Meehl, 1959; Goldberg, 1970; Dawes and Corrigan, 1974). This research suggested that nonlinear rules are not useful predictors in this task, since they do not perform as well as simple linear combinations of the scales. In particular, Goldberg (1965, 1972) suggested that a simple linear combination of five scales ($L + P_a + S_c - H_y - P_t$) outperforms any of the commonly used nonlinear rules or combination of these rules. So far, these results have not been refuted by any published research (e.g., Goldberg, 1969; Giannetti, Johnson, Klinger, and Williams, 1978). Furthermore during the years since the publication of his original research, there have also been quite a few unpublished attempts to "beat" the Goldberg formula using configural rules; none of them have been successful (Goldberg, personal communication).

I thank Dr. Lewis R. Goldberg for his help, both in supplying the data and in answering many of my questions. I also thank Robyn Dawes, Nimrod Kleinman, Avi Kluger, David Krantz, Dany Mingelgrin, and Yaacov Schul for their comments on earlier versions of this paper.

Address correspondence and reprint requests to Yoav Ganzach, Faculty of Management, Tel Aviv University, Tel Aviv, Israel, 69978. E-mail: yoavgn@post.tau.ac.il.

0749-5978/98 $25.00

Although comparison against the benchmark of simple linear rules is important in determining the usefulness of any possible nonlinear rule, the real test for these rules is whether they have *incremental validity* over linear rules; that is, whether they can explain some of the criterion variance which is not explained by the linear rule. This is the approach guiding the current research (see also Goldberg, 1965).

There is already evidence suggesting that nonlinear rules may have incremental validity. This evidence comes primarily from studies of nonlinearity in clinical *judgment*. Using the MMPI scales as independent variables, Goldberg (1970; see also Einhorn, 1974) found nonlinear accuracy in MMPI-based judgments of the likelihood of psychosis versus neurosis. He found that the residuals of a linear model of the *criterion* (the "true pathology") were correlated with the residuals of the linear model of the *judgment* (i.e., the judgment of pathology based solely on the MMPI). These results indicate that in making MMPI based clinical judgments, clinicians use incrementally valid nonlinear rules. However, Goldberg's (1970) study fell short of identifying these nonlinear rules. In fact, in a subsequent study, Goldberg (1971) found that out of the 65 nonlinear rules of the MMPI usually used by clinicians, only one was significantly correlated ($p < .05$) with the residuals of a linear model of the criterion. Since this is what would be expected by chance, this finding is, if anything, evidence against the incremental validity of the commonly used nonlinear rules.

Nevertheless, the results of Goldberg's (1970) and Einhorn's (1974) studies suggest that by studying clinical judgment it may be possible to discover incremtally valid nonlinear rules for the criterion. To do this, one needs to identify nonlinear elements which describe the residuals of the linear model of the judgment and examine if these elements are valid predictors of the residuals of the linear model of the criterion. Thus, in this paper I examine whether nonlinear rules of MMPI-based clinical judgments are useful in predicting the criterion—the true state of psychiatric patients. This approach may appear to put the carriage before the horse, since normally the rules which relate predictors to judgment ought to be derived from the rules which relate predictors to criterion and not vice versa. However, since previous research did find incrementally valid nonlinear rules for the judgment (Ganzach, 1995), but did not find such rules for the criterion, and since there is evidence of nonlinear accuracy in the judgment, this could be a possible approach.

Nonlinear Rules as Predictors of the Judgment and the Criterion

Ganzach (1995) identified nonlinear rules underlying the judgment of the likelihood of psychosis from the MMPI, which have incremental validity over a linear model of the judgment. Two of these rules are the focus of the current paper—they are used to achieve incremental validity in the criterion. One rule is associated with the integration of the neurotic scales of the MMPI, the scales originally derived from neurotic patients, and the other is associated with the integration of the psychotic scales, the scales which, in the current data, are related to psychotic disorders (see below for the classification of the scales in

the current study)[1]. The two rules have the same structure, and both are defined by the equation

$$SF = \left[\sum_{k=1}^{q} (X_k - F)^2 \right]^{1/2}, \tag{3}$$

where SF is the scatter of the psychotic [neurotic] scales, F is the mean of the psychotic [neurotic] scales, q is the number of the psychotic [neurotic] scales, and the X_ks are the scale values of the psychotic [neurotic] scales. These rules suggest that the scatter, or internal, variability of the neurotic and psychotic scales has incremental validity over their linear combination in predicting the pathology of a psychiatric patient (see Cronbach and Gleser, 1953, for a discussion of the concept of scatter in multiattribute profiles).

To see why scatter may be predictive of pathology, consider two profiles, equal on all scales except for two *psychotic* scales which have the same mean but differ in scatter. Whereas in the first profile (the low-scatter profile) the scores of the two scales are moderate, in the second profile (the high-scatter profile) one scale has a high score and the other has a low score. If the high-score scale is more indicative of the likelihood of psychosis than the low-score scale, the high-scatter profile is more likely to be of a psychotic patient than the low-scatter profile. In other words, in this case the higher the scatter of the psychotic scales, the higher the likelihood of psychosis[2]. Similarly, the neurotic scatter rule suggests that, other things being equal, the higher the scatter of the *neurotic* scales, the higher the likelihood of neurosis. Note that when a profile is either of a psychotic patient or of a neurotic patient, this second rule implies that the higher the scatter of the neurotic scales, the *lower* the likelihood of psychosis.

Finally, note that the two scatter rules are consistent with the fact that the relevant scales of the MMPI are related to the probability that a patient has a specific psychiatric syndrome. Since in the psychiatric classification system which is relevant to the current study, where the presence of even one strong psychotic [neurotic] symptom was a sufficient condition for a patient being

[1] Ganzach (1995) examined an additional rule, labeled the simple scatter rule. Whereas the two rules described in the text suggest a preliminary intradimension integration of the cues followed by dimension integration, the simple scatter rule suggests a one stage integration process. Although this latter rule has an incremental validity over a linear model of the judgment (though not over a linear model of the criterion, see Ganzach, 1997) this incremental validity is smaller than the incremental validity which is achieved by using multiple scatters. The reason is the nonlinear relationships associated with the psychotic dimension are opposite to those associated with the neurotic dimension (Ganzach, 1995).

[2] This relationship between scatter and pathology is labeled a disjunctive relationship. In principle the relationship between scatter and pathology may also be conjunctive—the higher the scatter, the lower the likelihood of pathology (see, for example, Einhorn, 1970 for further discussion of these two relationships). Such a relationship, which is associated with a case in which a low-score scale is more indicative of the likelihood of psychosis than a low-score scale, would be inconsistent both with the judgment of pathology (Ganzach, 1995) and with the psychiatric classification system (see the discussion below).

diagnosed as psychotic [neurotic], rules that assign higher weight to higher-value scales among the psychotic [neurotic] scales may be valid rules. For example, in this classification system, the presence of a manic symptom was sufficient for a diagnosis of psychosis, even if other psychotic symptoms such as paranoid or schizophrenic symptoms were missing. Therefore, when integrating the values of the manic and schizophrenic scales of the MMPI, assigning higher weight to the higher scale of the two may be justified.

THE DATA

The important studies that led to the bleak conclusions about the validity of nonlinear rules were based on data collected by Meehl (1959) and by Meehl and Dahlstrom (1960). These data included 1263 MMPI profiles of psychiatric patients—the patients' scores on the 11 most commonly used scales of the MMPI. The data also included the criterion—the diagnosis given to the patient in the clinic in which he/she received treatment. Forty-seven percent of the patients were diagnosed as psychotics and 53% were diagnosed as neurotics. These diagnoses were based primarily on information about the patient's past and present behavior which was collected by the clinic's staff, and, to a certain extent, on the results of various psychological tests. For some of the patients, the MMPI test results were available when the criterion was determined (i.e., when the clinic's diagnosis was made), but for other patients, the MMPI information was not available (Meehl, 1959). The profiles of the former patients are labeled "contaminated profiles," and the profiles of the latter patients are labeled "uncontaminated profiles" (see Table 1 for the number of patients in each group).

The data also included the results of an experiment conducted by Meehl in which 861 of the profiles (the profiles collected by Meehl, 1959) were given to 29 clinicians who were asked to judge each profile on an 11-step forced normal distribution scale from least psychotic (1) to most psychotic (11). The clinicians were instructed that the patients could be either psychotics or neurotics (see Meehl, 1959, for a detailed description of the experiment). Although these judgments are not directly relevant to the question of the validity of nonlinear

TABLE 1

Partial Correlations between the Psychotic and Neurotic Scatter and the Residuals of the Best Linear Model of the Criterion

	Meehl and Dahlstrom' Data	Meehl's Data			
		All profiles (n = 861)	Uncontaminated (n = 300)	Contaminated (n = 469)	Unknown (n = 92)
Psychotic scatter	.17***	.14***	.12*	.12**	.29**
Neurotic scatter	−.04	−.07*	−.16**	.00	−.09

*$p < .05$.
**$p < .01$.
***$p < .001$.

rules in predicting the criterion, they are relevant to the current paper, since our main a priori argument for the validity of these rules in predicting the criterion is Goldberg's (1965) and Einhorn's (1974) results concerning the nonlinear accuracy in the clinicians' judgments.

One aspect of the data which is important to the current study is that the MMPI scales of the 861 profiles have a clear dimensional organization (see Ganzach, 1995, for the results of a factor analysis of the scales). One dimension is associated with the neurotic scales of the MMPI (H_s, D, H_y, and P_t), another with the psychotic scales (F, P_d, P_a, S_c and M_a), and a third with scales that identify defensiveness in test taking (L and K). These dimensions, and in particular the neurotic and the psychotic dimensions, were likely to have played an important role in the process by which the clinicians made their judgments in Meehl's experiment (Ganzach, 1995). Note that while the scales F and P_d are usually not regarded as psychotic scales, in the current data they loaded heavily on the same factor as the scales originally derived from psychotic patients, most likely because the sample consisted only of psychotic and neurotic patients (Ganzach, 1995). Table 2 presents the intercorrelations among the scales, the criterion, and the two scatters as well as their means and standard deviations.

Finally, the relevance of these 40-year-old data could be questioned on the ground that both the diagnostic criteria and the test used to predict these criteria have been changed over the years. However, in my view, the timeliness of the data is not relevant to the question asked in the paper, and the demonstration of nonlinear relationships in decision making must not be limited to current tests or current diagnostic criteria. Furthermore, with regard to the question asked in the paper, there is no other set of data which has been studied as

TABLE 2

Means, Standard Deviations, and Intercorrelation for the MMPI Scales, the Criterion, and the Two Scatters

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	M	SD
1. Criterion	—													.47	.50
2. Psychotic scatter	.17	—												10.0	4.0
3. Neurotic scatter	−.09	.21	—											9.5	4.7
4. L scale	.13	.02	−.03	—										5.06	6.1
5. F scale	.18	.38	.19	−0.4	—									58.7	8.9
6. K scale	.07	−.04	−.25	.37	−.36	—								53.2	9.2
7. H_s scale	−.15	.30	.04	.11	.15	.01	—							67.3	16.1
8. D scale	−.13	.43	.46	.00	.31	−.29	.53	—						75.2	17.6
9. H_Y scale	−.18	.34	.02	.11	.12	.08	.81	.58	—					67.3	12.2
10. Pd scale	.15	.48	.08	−.02	.38	.04	.15	.28	.26	—				66.7	11.9
11. Pa scale	.17	.37	−.17	.00	.49	−.21	.32	.43	.38	.44	—			60.7	11.3
12. Pt scale	−.03	.56	.36	.06	.41	−.22	.47	.75	.54	.40	.56	—		71.2	16.2
13. Sc scale	.19	.66	.20	.01	.65	−.14	.43	.31	.45	.49	.63	.77	—	69.6	16.2
14. Ma scale	.15	−.01	−.10	−.09	.34	−.20	.08	−.13	.03	.30	.23	.10	.33	57.3	12.0

extensively as Meehl's data, and the detection of nonlinear relationships in these data is in many respects a benchmark for a convincing demonstration of the validity of nonlinear rules.

RESULTS

The incremental validity of the scatter rules: In the following analysis I adhere to Goldberg's (1965) approach and examine the incremental validity of the scatter rules by comparing the cross validity of the best linear model (Goldberg's five-scales model) to the cross validity of a model that includes, in addition to the five terms of the best linear model, also the two scatter rules.

Following Goldberg (1965), I have built a regression model using Meehl and Dahlstrom's (1960) data and have used the parameters of this model to predict the probability of psychosis in Meehl's data. The correlation between this predicted probability and the criterion is .44 for the five-scale linear model (see Goldberg, 1965) and .46 for a model that includes the five scales and the two scatter rules. In addition, when the regression model is built using Meehl's data and the validation is performed on Meehl and Dahlstrom's data, the correlation between the predicted probability and the criterion is .41 for the best linear model and .45 for a model that includes also the two scatter rules. Thus, a combination rule that adds the two scatters to the best linear combination rule outperforms Goldberg's best linear combination rule.

Another way to examine the incremental validity of the two scatter rules is to correlate, across profiles, the residuals of the best linear model and the values of the two scatters. These correlations are presented in column 2 of Table 1 for Meehl and Dahlstrom's data and in column 3 for Meehl's data. Of the four correlations, three are significant, and the signs of the correlations are consistent with the notion that among the psychotic [neurotic] scales, the highest psychotic [neurotic] scales are the most important indicators of the probability of psychosis [neurosis].

Finally, note that to be consistent with Goldberg's (1965) method, the previous analyses did not use logistic regression, which may be a more appropriate method for a dichotomous dependent variable. However, when the classification base rate is about 50% and the R^2 is low, the difference between the two methods is very small. Indeed, when the significance levels in Table 1 were assessed by a logistic regression that examined the marginal contribution of the scatter rules in predicting the criterion, the results of were similar to the results reported in the table.

The Effect of Contamination: It could be argued that, because of criterion contamination, the validity of the nonlinear rules reflects nonlinearity in the MMPI integration rules of the clinicians that produced the clinic's diagnosis rather than nonlinearity in the "true" criterion. To examine this possibility, the partial correlations between the two scatters and the criterion were calculated separately for the 300 uncontaminated profiles and the 461 contaminated profiles in Meehl's data (there was no information regarding the degree of

contamination with regard to the Meehl and Dahlstrom data or with regard to 92 of the profiles in Meehl's data). These partial correlations are also presented in Table 1. If anything, these results suggest that the nonlinear rules are more valid for the uncontaminated profiles than for the contaminated profiles.

The Determinants of Nonlinear Accuracy: Support for the notion that the nonlinear accuracy found by Goldberg (1965) and Einhorn (1974) is explained by the scatter rules is obtained by correlating the nonlinear accuracy of the judges with the weights the two scatters in their judgments. In the analysis below, nonlinear accuracy is operationalized as the correlation between the residuals of a linear model of the judgments and the residuals of a linear model of the criterion using the 11 MMPI scales as independent variables (the entire set of 11 scales was used since, with regard to the judgment, there was no a priori optimal subset of the scales); the weight of the neurotic scatter in the judgment is operationalized as the partial correlation between the clinician's judgments and the neurotic scatter, controlling for the linear effect of the 11 scales, and the weight of the psychotic scatter in the judgment is operationalized as the partial correlation between the judgments and the psychotic scatter, controlling for the linear effects of the 11 scales. (Since the correlation between the two scatters in the stimuli was small [$r = -.16$, $p > .4$], controlling for them has very little impact on the results).

The data suggest that, to a large extent, nonlinear accuracy is explained by the weights of the two scatters in the judgment. The correlation between the weight of the neurotic scatter in the judgment and accuracy is $-.33$ ($p < .1$), the correlation between the weight of the psychotic scatter in the judgment and accuracy is .57 ($p < .001$), and the multiple correlation between the two weights and accuracy is .62. These correlations are consistent with the notion that heavier reliance on the scatter rules leads to higher nonlinear accuracy. Note also that the signs of these correlations are consistent with the notion that greater accuracy is achieved by assigning heavier weight to the highest psychotic [neurotic] scale.

DISCUSSION

In his 1965 paper, Goldberg examined the validity of a large number of nonlinear models as predictors of the criterion in Meehl's data, but none of them performed as well as simple linear combination rules. Subsequently, in a paper labeled "The search for configural relationships in personality assessment: The diagnosis of psychosis vs. neurosis from the MMPI," Goldberg (1969) reported the results of additional attempts to find such models, but again, none of the trial models did better than a simple linear combination of five scales. These results led Goldberg, and later on Dawes and Corrigan (1974), to emphasize the robustness of linear models as predictors of real-life outcomes.

With this in mind, it is interesting that, in the same data that led to the conclusions about their poor predictive power, nonlinear rules have incremental validity over the best linear combination role, and that they are good predictors

of judges' nonlinear accuracy. It also should be noted that these rules are not the result of a fishing expedition, but are derived from the analysis of clinical judgment, and that they cannot be explained by criterion contamination, since contamination is, if anything, detrimental to their validity.

Are these results generalizable? That is, do they indicate that incrementally valid nonlinear rules could be found in domains other than the diagnosis of psychosis versus neurosis from the MMPI? The current data do not provide an answer to this question. On the one hand, it is possible that the nonlinear relationships between the MMPI scales and pathology are a unique example of nonlinear relationships between predictors and criteria stemming from the fact that even one strong psychotic [neurotic] characteristic is sufficient for the diagnosis of the patient as psychotic [neurotic]. Furthermore, it is also possible that attempting to identify nonlinear relationships between predictors and criteria on the basis of nonlinear relationships between predictions and judgments is unwarranted because judgments are often characterized by erroneous configural strategies (Ganzach, 1993; 1997). The third nonlinear rule which was found in the judgments of Meehl's experiment—the rule associated with the integration of the neurotic and psychotic scales (Ganzach, 1995)—is a relevant example, since it characterizes the judgment but not the criterion.

On the other hand, nonlinear relationships between predictors and criteria in which extreme predictors are more indicative of the criterion than moderate predictors may be possible. For example, success in professional football may be determined by the player's best skill (e.g., kicking, throwing) rather than by the average of his skills, whereas success in obstacle-course running may be determined by the runner's worst skill (jumping, running). It is also possible that nonlinear characteristics are introduced into the relationships between predictors and criteria because many real-world criteria include strong judgmental elements. The opinion of a supervisor is used as a measure of employee performance, and the impression of a clinician is an important input into diagnostic decisions. Therefore, since nonlinearity is a strong characteristic of the judgment (e.g., Camerer, 1981; Ganzach, 1994), it may also characterize the criterion.

Dawes and Corrigan (1974) argued convincingly that, because of the robustness of linear models, when the relationships between predictors and criterion are conditionally monotone, even if the true model is nonlinear, the incremental variance of nonlinear elements will be small. However, this does not mean that the identification of nonlinear relationships between predictors and criteria is unimportant or futile. First, with the advent of the computer age, the accumulation of large databases is likely to supply researchers with more and more power to identify stable nonlinear relationships; second, when the predictions of the model are important—for example, when they are used to make decisions about a large population—even a small increase in explained variance due to nonlinear relationships may be associated with a large incremental value; and third, the identification of nonlinear relationships may very well supply insights into the causal processes that relate predictors to outcome.

REFERENCES

Camerer, C. F. (1981). General conditions for the success of bootstrapping models. *Organizational Behavior and Human Performance,* **27,** 412–422.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin,* **50,** 546–473.

Dawes, R. M., and Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin,* **81,** 95–106.

Einhorn, H. J. (1970). The use of nonlinear noncompensatory models in decision making. *Psychological Bulletin,* **73,** 221–230.

Einhorn, H. J. (1974). Cue definition and residual judgment. *Organizational Behavior and Human Performance,* **12,** 30–49.

Ganzach, Y. (1993). Goals as determinants of nonlinear, noncompensatory judgment strategies. *Organizational Behavior and Human Performance,* **56,** 422–440.

Ganzach, Y. (1994). Theory and configurality in expert and layman judgment. *Journal of Applied Psychology,* **79,** 439–448.

Ganzach, Y. (1995). Nonlinear models of clinical judgment: Meehl's data revisited. *Psychological Bulletin,* **118,** 422–429.

Ganzach, Y. (1997). Theory and configurality in clinical judgment of expert and novice psychologists. *Journal of Applied Psychology,* **82,** 954–960.

Giannetti, R. A., Johnson, J. H., Klinger, E. E., & Williams, T. A. (1978). Comparison of linear and nonlinear MMPI diagnostic methods with an uncontaminated criterion. *Journal of Consulting and Clinical Psychology,* **46,** 1046–1052.

Goldberg, L. R. (1965). Diagnosticians vs. diagnostic signs: The diagnosis of psychosis vs. neurosis from the MMPI. *Psychological Monographs,* **79.**

Goldberg, L. R. (1969). The search for configural relationships in personality assessment: The diagnosis of psychosis vs. neurosis from the MMPI. *Multivariate Behavioral Research,* **4,** 523–536.

Goldberg, L. R. (1970). Man versus model of man: A rationale plus some evidence for improving clinical inferences. *Psychological Bulletin,* **73,** 422–434.

Goldberg, L. R. (1971). Five models of clinical judgment: An empirical comparison between linear and nonlinear representation of the human inference process. *Organizational Behavior and Human Performance,* **6,** 458–479.

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

Meehl, P. (1959). A comparison of clinicians with five statistical methods of identifying psychotic MMPI profiles. *Journal of Counseling Psychology,* **6,** 102–109.

Meehl, P., and Dahlstrom, W. G. (1960). Objective configural rules for discriminating psychotic from neurotic MMPI profiles. *Journal of Consulting Psychology,* **24,** 375–387.

Received: July 4, 1997